

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using bw model

Run on: February 26, 2005, 17:00:07 ; Search time 2362 Seconds
(without alignments)
10216.214 Million cell updates/sec

Title: US-10-621-901-26
Perfect score: 498
Sequence: 1 ggccctgacaaagatgcct.....aaattctgcacagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 470823 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

1:  gb_ba:*
2:  gb_hgt:*
3:  gb_in:*
4:  gb_cm:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_um:*
13: gb_un:*
14: gb_vi:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	189.4	38.0	3015	6	C0578439	C0578439 Sequence
2	189.4	38.0	3037	3	AY044504	AY044504 Drosophila
3	125.4	25.2	6552	6	C0578438	C0578438 Sequence
4	125.4	25.2	71504	3	AC002473	AC002473 Drosophila
5	125.4	25.2	88933	2	AC017870	AC017870 Drosophila
6	125.4	25.2	171594	3	AC099019	AC099019 Drosophila
7	125.4	25.2	288451	3	AE003456	AE003456 Drosophila
8	84.8	17.0	250968	2	AC014140	AC014140 Drosophila
9	83.2	16.7	2486	6	C0583587	C0583587 Sequence
10	83.2	16.7	6105	6	C0583586	C0583586 Sequence
11	83.2	16.7	154772	3	AC0098214	AC0098214 Drosophila
12	83.2	16.7	165267	3	AC005711	AC005711 Drosophila
13	83.2	16.7	246230	3	AE003636	AE003636 Drosophila
14	72.4	14.5	5349	6	C0578543	C0578543 Sequence
15	65.8	13.2	2184	6	C0578538	C0578538 Sequence
16	65.8	13.2	4399	6	C0578537	C0578537 Sequence
17	59	11.8	2552	2	C0578544	C0578544 Sequence
18	43.6	8.8	17530	2	CR352298	CR352298 Dario ret
19	43.6	8.8	181320	5	BX547992	BX547992 Zebrafish

20	42.2	8.5	222994	2	AC094289	AC094289 Rattus no
21	41.8	8.4	2175	9	AF401643	AF401643 Homo sapi
22	41.8	8.4	3258	9	AY273896	AY273896 Homo sapi
23	41.8	8.4	5054	6	CO851267	CO851267 Sequence
24	41.8	8.4	5054	6	AK128854	AK128854 Homo sapi
25	41.8	8.4	5238	9	HSMB07493	BS647349 Homo sapi
26	41.8	8.2	3998	5	AJ79335	AJ79335 Gallus ga
27	40.8	8.2	185456	2	AC091379	AC091379 Mus muscu
28	40.8	8.2	217530	10	AC121845	AC121845 Mus musc
29	40.4	8.1	2594	10	AF203701	AF203701 Mus musc
30	40.4	8.1	2675	10	BC072114	BC072114 Mus musc
31	40.4	8.1	2798	10	AF213260	AF213260 Mus musc
32	40.4	8.1	108609	2	AY379774	AY379774 Medicago
33	40.4	8.1	110000	8	CR3812127_20	CR3812127_20
34	40.4	8.1	158372	2	BX47014	BX47014 Danio rer
35	40.4	8.1	266607	2	BX571847	BX571847 Danio rer
36	40.2	8.1	8866	6	AX277977	AX277977 Sequence
37	40.2	8.1	8866	6	AX233666	AX233666 Sequence
38	40.2	8.1	16771	3	CEY32F6B	AL021475 Caenorhab
39	40.2	8.0	42789	3	U39893	U39893 Acanthosabi
40	40	8.0	67575	2	AC098997	AC098997 Mus muscu
41	39.6	8.0	5882	6	AX345447	AX345447 Sequence
42	39.6	8.0	145695	2	AC015690	AC015690 Homo sapi
43	39.6	8.0	151939	9	AP001980	AP001980 Homo sapi
44	39.6	8.0	167960	2	AC007559	AC007559 Homo sapi
45	39.6	8.0	169801	2	AC018429	AC018429 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	3015 bp	DNA	linear	PAT 02-FEB-2006
CO578439	DEFINITION	Sequence 6197 from Patent WO0171042.			
CO578439	ACCESSION	CO578439			
CO578439.1	VERSION	GI:41640939			
ORGANISM	SOURCE	Drosophila sp.			
		Drosophila sp.			
		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	AUTHORS	Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.			
TITLE		Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof			
JOURNAL		Patent: WO 0171042-A 6197 27-SEP-2001;			
		PB Corporation (NY) (US)			
FEATURES	Location/Qualifiers				
source	1..3015	/organism="Drosophila sp."			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:7242"			
ORIGIN					
Query Match	38.0%;	Score 189.4;	DB 6;	Length 3015;	
Best Local Similarity	61.3%;	Pred. NO. 1.8e-42;			
Matches 304;	Conservative 0;	Mismatches 192;	Indels 0;	Gaps 0;	
Oy	1	GGCCCTGAGACAGATGCGCTTCGACATCCACCAAGTAATGAGCGTCTCAAGATCAAAAT	60		
Db	1127	GGAGCCGGCGGAGATGCACCTTGCTCTGACCTCGAGTTCGGCGGAGTCCCGATGAAAT	1166		
Oy	61	GCCACCAATGGAAGTCTTAAATAACAGAGAGAAAACTTTGGCACACTGATGGGAA	120		
Db	1187	GCCCAATGAGGGGTATATCAGAGACACGATCCAAAGCCCTATCCGCTGAAATGAGGA	1246		
Oy	121	GGAATTGAATGCAAGAACAGAAAGGAGGAGCATGCACCTCAATTAATACTATTGCTGCC	180		
Db	1247	GGAGCGGAGTCCGAGTTCGTTGTAAGGAAACTTTGGCCGCCCAACTTTTGCTTGGTAGCG	1306		
Oy	181	CAATTAATTTCTGAGTCCGTGATCTTAACTACTACCTTAGGTGATATCTAATATGAC	240		

Db 1307 CAGTTATATCTGGAATGAGATGCTATCTACCGTTGGAGGTCCTACATGAC 1366
QY 241 GATTAACACCAAAAAATCTAAAAACACGACGACTGATTAATTTCTTATTTCTTCGATG 300
Db 1367 GATTAATCTAAAAATCTAAAAACCTCCGGCTTTGTGATCTTTCATCTTTTGGCATG 1426
QY 301 CTCGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTTCAATTTGCCA 360
Db 1427 TTGGGCTCTGCTCCATTTGGATACGCCCTGGCATCTTCTGCGCCCTGATATCGACCA 1486
QY 361 ACCTTGACCGCAACATTTGATTAATATGATCTAGATGTTATGAGATGCTGTTAAAT 420
Db 1487 CAATATCATCTCGGTGATCAACAACAAGATCCCCGCTGGGCTGGGCTGTGCTGGGC 1546
QY 421 TGGGATATCTAGATCAACTTAAATTTTGGCAAGTTGATTTGATTTCTTAA 480
Db 1547 TGGCTGTATATGGCGCGCTGCTCTCTTCTCGGAGTTTCTCTCCATGTTTCCCAA 1606
QY 481 ATTCTGCCAAGAGCTG 496
Db 1607 GAATGCCCAAGACAG 1622

RESULT 2
AY047504 3037 bp mRNA linear INV 16-AUG-2001
LOCUS Drosophila melanogaster GH01304 full length cDNA.
DEFINITION AY047504
ACCESSION AY047504.1 GI:15010375
KEYWORDS FLI_CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Bukayoka; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 3037)

REFERENCE
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phoumanavong,S., Wan,K.,
Yu,C., Lewis,S.E., Rubin,G.M. and Celinker,S.
Direct Submission
Submitted (19-JUL-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

TITLE
JOURNAL
COMMENT
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNA, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
source location/Qualifiers

1..3037
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="58D1-58D1"
/clone="GH01304"
1..3037
gene

CDS
/gene="CG3380"
/note="alignment with genomic scaffold AB003456"
/db_xref="FLYBASE:FBgn0034716"
800..2896
/gene="CG3380"
/note="longest ORF"
/odon_start=1
/product="GH01304P"
/protein_id="AAK7236.1"
/db_xref="GI:15010376"
/translation="MPASGYSRERFTDLPILKLTLLYGIYICISMTYAVFNCTI
TTIBKRFKPSKNTGILISGNDISQTVSAVALAYAKGHRPRMIGFGLTIVFCIL
TAPHPFLVGPEDALALISEFGMPDENATMEALIEORSKTIICRLNKGABCEVGEIN
FAPOLLFYAOFISGIGSLYYTLGVSYMDNDTKSKTPALSLISYFLRMLGPAIGYA
LASFCILRIAPOMHPVINNKDPRMIGAMLMVLWGLISFSGYLSMFKPELPRAY
ARRKYENRREKERLSVSKTERLALDOKTTEAASFODMKCTRRALITNKTY
MONTLSSIFYLYGTPYMIFFPKYIEVORSAASMTGTVALAFSAVGLISGFI
ISRYRPARAYMAWVYIVGFLVAGILATAFICRPNESSVIVNIHDSLAGNTTCN
SACSQDYRYSFVCGENNTYISACHAGAKLIVNSBKRTIFDSCIPSDDGNST
QPKRLTSPDLSNDEVSONTSVHSOLEALNAGAMPACVNCMTQFVAELAWCCLEF
VGASGRASNFIVSRCPKEDKTAAMFGMTLCSMIAFIIPSDIFFGMDVDRVCLWVG
TCNKGNCWLYDPLASRYTLNFTAAVFIAGAFIDLVGVYVAKDLKIFDEYVEVEMK
IVQHEBEANENKNTET"

ORIGIN
Query Match 38.0%; Score 189.4; DB 3; Length 3037;
Best Local Similarity 61.3%; Pred. No. 1,8e-42;
Matches 304; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 1 GGCCCTGGAACAAGATCCCTTGACCTCACCAGAAATATGAGCTGTACCAAT 60
Db 1130 GGACCCGGGAGAGATCACTTGCTGACCTCGGAGTTGGCGGAGTCCGATGA 1189
QY 61 GCCACATTAAGATTCTTAATAAAGAGAAACCTTGGCCACATGATGCGAA 120
Db 1190 GCCACATTAAGAGCTATCAAGAGAGAGATCCAAACCTATGCCCTTGAATGA 1249
QY 121 GGAATTGAATGCGAAACAGAGAGAGATCGACCTCAATAATATCTATTTGCTGC 180
Db 1250 GGAAGGAGAGGAGAGTGGTGAAGAACTTGGCCCACTTTGCTTGGTGAAG 1309
QY 181 CAATTAATTTCTGAGTGGTGAATTTTATCTACATCTTATGATCTTATGAGAC 240
Db 1310 CAGTTATATCTGGAATGAGATGCTATCTACACCTTGGAGATGTCCTACATGAG 1369
QY 241 GATTAACACCAAAAAATCTAAAAACACGACGCTATTAATTTCTTATTTCTTCGATG 300
Db 1370 GATTAATCTAAAAATCTAAAAACCTCCGGCTTTGTAGCTTTCATCTTTTGGCATG 1429
QY 301 CTCGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTTCAATTTGCCA 360
Db 1430 TTGGGCTCTGCTCCATTTGGATACGCCCTGGCATCTTCTGCTGGGCTGATATG 1489
QY 361 ACCTTGACCGCAACATTTGATTAATATGATCTAGATGTTATGAGATGCTGTTAAAT 420
Db 1490 CAATATCATCTCGGTGATCAACAACAAGATCCCCGCTGGGCTGGGCTGTGCTGGGC 1549
QY 421 TGGGATATCTAGATCAACTTAAATTTTGGCAAGTTGATTTGATTTCTTAA 480
Db 1550 TGGCTGTATATGGCGCGCTGCTCTCTTCTCGGAGTTTCTCTCCATGTTTCCCAA 1609
QY 481 ATTCTGCCAAGAGCTG 496
Db 1610 GAATGCCCAAGACAG 1625

RESULT 3
COS78438/c 6522 bp DNA linear PAT 02-FEB-2004
LOCUS COS78438
DEFINITION Sequence 6196 from Patent W00171042.
ACCESSION COS78438
VERSION COS78438.1 GI:41640938

KEYWORDS	Drosophila sp.
SOURCE	Drosophila sp.
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	Venter, J.C., Adams, M., Li, P.W. and Myers, E.W. Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof Patent: WO 0171042-A 6196 27-SEP-2001; PE Corporation (NY) (US)
AUTHORS	1
TITLE	
JOURNAL	
FEATURES	Location/Qualifiers
SOURCE	1..6522 /organism="Drosophila sp." /mol_type="unassigned DNA" /db_xref="taxon:7242"
ORIGIN	
Query Match	25.2%; Score 125.4; DB 6; Length 6522;
Best Local Similarity	65.6%; Pred. No. 2.4e-24;
Matches 183; Conservative	0; Matches 96; Indels 0; Gaps 0;
Oy	1 GGCCCTGACAGAAATGCCCTTGCACTCACCGAATTATGGAGCTGTACAGTCAAAAT 60
Db	3510 GGACCCGGCGAGATGCATTGCTTTGACTTCGGAGTTTCGGGGGATCCCCGATGAATAAT 120
Oy	61 GCCACCATTTGAAGTTCTTAATTAACAGAAAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db	3450 GCCACAATAGAGGCTATGAGAGAGAGAGATCACAACCCTATGCGCTTGAATGGAGAGA 3391
Oy	121 GGAGTTGAATGCGAAACAGAAAGGAGCATCGCACTCAAATTAATCTATTGTCGCC 180
Db	3390 GGAGGGGAGTCCGAGTGGTGAAGGAAACTTTGGCCCCCACTTTCCTCTTGGTAGCG 3331
Oy	181 CAATTAATTTCCGAGTGGTGGATCTTTACTACACTTGAAGGTATCTCATATATGGAC 240
Db	3330 CAGTTCATATCTCGAATGAGAGATGCTATATCACTGAGTTGGGAGTGTCTCAATGAC 3271
Oy	241 GATTAACACAAAAAATCTTAACACACGACACTGATTAAT 279
Db	3270 GATTAATCAAAAAAAAAATCCAAAATCTCGGCTTTGTTGAACT 3232
RESULT 4	
AC002473/c	71504 bp DNA linear INV 20-JAN-1998
LOCUS	Ac002473
DEFINITION	Drosophila melanogaster (PI D502833 (D61)) DNA sequence, complete
ACCESSION	AC002473 L81545 L81544 L81543 L81542 L81541 L81540 AC000798
KEYWORDS	Sequence.
ORGANISM	Ac000801 L81539 L81538 L81537 AC000797 AC000796 AC000802 AC000800
KEYWORDS	Ac000799 AC0001823 AC0001826 AC0001824 AC001822 AC001337 AC001338
KEYWORDS	Ac001825 AC001339
KEYWORDS	Ac002473.1 GI:2337890
KEYWORDS	HTE.
KEYWORDS	Drosophila melanogaster (fruit fly)
KEYWORDS	Drosophila melanogaster
KEYWORDS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 71504) Ceiniger, S.E., Aghavan, A., Arcaina, T.T., Baxter, E., Doyle, C.M., Fafan, D.E., Flanagan, J., Houston, K.A., Hummachi, S.R., Karre, K., Kearney, L., Kim, S.H., Ko, C.L., Li, M., Lomocan, M.A., Mazda, P., Mok, M.S., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Santoe, R.F., Snit, E., Stevo, V., Subramanian, S., Towne, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.V. and Kimmel, B.
TITLE	Sequencing of Drosophila chromosome 2R, region 5B2
JOURNAL	Unpublished (1997)
REFERENCE	2 (bases 1 to 71504) Martin, C.H., Alcivar, D.A., Arcaina, T.T., Bondoc, M.M., Chiang, A., Ciliz, P.A., Davis, C.A., Doyle, C.M., Ericsson, C.L., Fafan, D.E.,

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (20-AUG-1997) Berkeley Drosophila Genome Project, MS 74-157, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.hgc.lbl.gov/sequence-archive.html>) or send email to drosophila@genome.lbl.gov.
 Library location: 30-49.
 This PI was assembled from the following subclones: 2_a4 (L81545), 1_f7 (L81544), 2_e9 (L81543), 2_b2 (L81542), 2_h6 (L81541), 2_b11 (L81540), 1_f9 (AC000798), 2_d6, 2_g7 (AC000801), 2_c11 (L81539), 1_f5 (L81538), 2_f5 (L81537), 1_e12 (AC000797), 1_b7, 2_a11, 1_f11, 1_c2 (AC000796), 2_g8 (AC000802), 2_f2 (AC000800), 2_c4 (AC000799), 2_c9 (AC001823), 2_f6 (AC001826), 2_e3 (AC001824), 2_a1 (AC001822), 1_c11 (AC001337), 2_f6 (AC001338), 2_g3 (AC001825), 2_f8 (AC001339).
 FEATURES
 source
 1. 71504
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /chromosome="2R"
 /map="58C2-58C7"
 /clone="p1 DS02833 (D61)"
 ORIGIN
 Query Match 25.2%; Score 125.4; DB 3; Length 71504;
 Best Local Similarity 65.6%; Pred. No. 2,5e-24;
 Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
 Oy 1 GGCCCTGGACAAGATGCGCTTCGACATCACCAGATATGAGAGCTGTACGATCAAAAT 60
 Db 61567 GGAGCCCGCGAGATGACATTCGCTCGACCTCGAGTGGGGGATGCCATGAATAAT 61508
 Oy 61 GCCACATTTGAGCTCTTAATAAGAGAAGAGAAACTTGTGCCACACTGATGGCGAA 120
 Db 61507 GCCACAAATGAGGCTATCGAGGACGACGATCCAAACCTTATGCCCTTGAATGAGGA 61448
 Oy 121 GGAATTGGAATGCGAAACAGAGAAGAGAGATGCACTCAATATATACTATTTGCTGCC 180
 Db 61447 GGAGCGGAGTGCAGAGTGGGTGAAGAACTTTGGCCCCCACTTTTGCTCTTGGTAGG 61388
 Oy 181 CAATTAATTTTGGAGTGGTGGATCTTTATACTACACTTTAAGGTATCTATATGAC 240
 Db 61387 CAGTTCAATCTATCGAATTTGAGGATGCTATACTACACCTTGGGAGTGTCTTACATGAC 61328
 Oy 241 GATTAACCCAAAATCTAATAACACAGACATGATAAGT 279
 Db 61327 GATTAATCTAATAAATCTCAAAACTCCGGCTTTGTTGAGT 61289
 RESULT 5
 AC017870/c
 LOCUS AC017870 88933 bp DNA linear HTG 09-DEC-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AC017870
 VERSION AC017870.1 GI:6553320
 KEYWORDS HTG; HTGS; PHASE2.
 SOURCE Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 88933)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT This sequence was identified as CDL10212423 by the submitter. For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1. 88933
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"

ORIGIN

Query Match 25.2%; Score 125.4; DB 2; Length 88933;
 Best Local Similarity 65.6%; Pred. No. 2.5e-24;
 Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 GGCCTTGACAAAGATGCCCTTTCAGCTCACCAAGATATGAGCTGTCTACATCAAAAT 60
 DB 64160 GGACCCGCGCAGAGATGACCTTGTCTGACCTCGGAGTCCGCGGAGATGCCCGATGAAAAAT 64101
 QY 61 GCCACCATTAAGTCTTAATAAACAAGAGAACTTTGCGCCACATGATGGGAA 120
 DB 64100 GCCACATGAGAGGCTATGAGAGAGACGATCCAAAGCCCTATGCGCCTGAATGGAGGA 64041
 QY 121 GGAATTGAATGCGAAACAGAAAGGAGCATGCGACCTCAATTAATATTGCTGCC 180
 DB 64040 GGAGCGGAGTGCAGAGTCCGATGGAAGAACTTTGCGGCCCAACTTTGCTTCTGTAGCG 63981
 QY 181 CAATTAATTTCTGAGTGGTGGATCTTTATATCACTTATAGTGTATCTATATGAC 240
 DB 63980 CAGTTCAATATCTGGAATTTGAGAGATCGCTATACACAGTTGGAGATGCTCTCATATGAC 63921
 QY 241 GATTAACCAAAAATCTTAACACGACGACTGATAGT 279
 DB 63920 GATTAATCTAAAAATCCAAAATCCGCGCTTTGTAGT 63882

RESULT 6
 AC099019 171594 bp DNA linear INV 08-NOV-2001
 LOCUS Drosophila melanogaster, chromosome 2R, region 58C-58D, BAC clone
 DEFINITION BACR02012, complete sequence.
 AC099019
 VERSION AC099019.1 GI:16798951
 SOURCE HTG.
 ORGANISM Drosophila melanogaster (fruit fly)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 171594)
 Celisner, S.E., Adams, M.D., Krommiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Branton, R.C., Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Chapple, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frise, E., Galle, R.F., Gary, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Scapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
 Sequencing of Drosophila chromosome, 2R, region 58C-58D

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 171594)
 AUTHORS Celisner, S.E., Adams, M.D., Krommiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Branton, R.C., Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Chapple, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frise, E., Galle, R.F., Gary, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Scapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

COMMENT
 TITLE Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
 JOURNAL Sequence submitted by: Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
 This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES
 source
 1. 171594
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /strain="Y; cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="2R"
 /map="58C-58D"
 /map="BACR02012 (D1312)"
 /clone_id="RPCT-98 (Roswell) Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACs.61"

ORIGIN

Query Match 25.2%; Score 125.4; DB 3; Length 171594;
 Best Local Similarity 65.6%; Pred. No. 2.5e-24;
 Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 GGCCTTGACAAAGATGCCCTTTCAGCTCACCAAGATATGAGCTGTCTACATCAAAAT 60
 DB 20134 GGACCCGCGCAGAGATGACCTTGTCTGACCTCGGAGTCCGCGGATGCCCGATGAAAAAT 20193
 QY 61 GCCACCATTAAGTCTTAATAAACAAGAGAACTTTGCGCCACATGATGGGAA 120
 DB 20194 GCCACATGAGAGCTATGAGAGAGAGACGATCCAAAGCCCTATGCGCCTGAATGGAGGA 20253
 QY 121 GGAATTGAATGCGAAACAGAAAGGAGCATGCGACCTCAATTAATATTGCTGCC 180
 DB 20254 GGAGCGGAGTGCAGAGTCCGATGGAAGAACTTTGCGGCCCAACTTTGCTTCTGTAGCG 20313
 QY 181 CAATTAATTTCTGAGTGGTGGATCTTTATATCACTTATAGTGTATCTATATGAC 240
 DB 20314 CAGTTCAATATCTGGAATTTGAGAGATGCTATATCACTTATAGTGGAGATGCTCATATGAC 20373
 QY 241 GATTAACCAAAAATCTTAACACGACGACTGATAGT 279
 DB 20374 GATTAATCTAAAAATCCAAAATCCGCGCTTTGTAGT 20412

RESULT 7
 AE003456/c 288451 bp DNA linear INV 22-MAR-2004
 LOCUS Drosophila melanogaster chromosome 2R, section 64 of 74 of the
 DEFINITION complete sequence.

ACCESSION	AE003456	AE002575	AE013599
VERSION	AE003456.3	GI:45445355	
KEYWORDS			
ORGANISM	Drosophila melanogaster (fruit fly)		
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	1 (bases 1 to 288451) Adams,M.D., Celinker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoeking,R.A., Gallie,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutcliffe,G., Wortman,J.R., Yeandle,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L., Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,A., An,H.J., Baxendale,A., Basley,E.M., Beeson,K.Y., Benos,P.V., Bernan,B.P., Bhandari,D., Bolshakov,S., Borkov,D., Botchan,M.R., Bouck,J., Brockstein,P., Brottler,P., Burdick,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Duncan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S., Fleischmann,W., Foster,C., Gabriellian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Hartie,N.V., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kienlen,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Laeko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Moritz,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Musy,D.M., Nelson,D.L., Nelson,K.A., Nixon,K., Nusser,D.R., Paclet,J.M., Palazolo,M., Peltman,G.S., Pan,S., Poland,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Sanders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svitskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weisenbach,J., Williams,S.M., Woodgett,J., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.B., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.		
TITLE	The genome sequence of Drosophila melanogaster		
JOURNAL	Science 287 (5461), 2185-2195 (2000)		
MEDLINE	20196006		
PUBMED	10731132		
REFERENCE	2 (bases 1 to 288451) Celinker,S.E., Wheeler,D.A., Kronmiller,B., Carlson,J.W., Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Fries,E., Hodgson,A., George,R.A., Hoeking,R.A., Laverly,T., Mazy,D.M., Nelson,C.R., Paclet,J.M., Park,S., Pfeiffer,B.D., Richards,S., Sodergren,E.J., Svitskas,R., Taber,P.E., Wan,K., Stapleton,M., Sutcliffe,G., Venter,C., Weinstock,G., Scherer,S.E., Myers,E.W., Gibbs,R.A. and Rubin,G.M.		
TITLE	Finishing a whole-genome shotgun: release 3 of the Drosophila melanogaster euchromatic genome sequence		
JOURNAL	Genome Biol. 3 (12), RESEARCH0079 (2002)		
MEDLINE	22426065		
PUBMED	12537568		
REFERENCE	3 (bases 1 to 288451) Matera,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S., Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E., Smith,C.D., Tupy,J.L., Whitfield,E.J., Beyrakeroglou,L., Bertram,B.P., Betencourt,B.R., Celinker,S., de Grey,A.D., Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J., Shu,S.O., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M., Rubin,G.M. and Lewis,S.E.		
TITLE	Annotation of the Drosophila melanogaster euchromatic genome: a systematic review		
JOURNAL	Genome Biol. 3 (12), RESEARCH0083 (2002)		
MEDLINE	22426069		
PUBMED	12537572		
REFERENCE	4 (bases 1 to 288451) Kaminker,J.S., Bergman,C.M., Kronmiller,B., Carlson,J., Svitskas,R., Patel,S., Fries,E., Wheeler,D.A., Lewis,S.E., Rubin,G.M., Ashburner,M. and Celinker,S.E.		
AUTHORS	The transposable elements of the Drosophila melanogaster euchromatic: a genomics perspective		
TITLE			
JOURNAL	Genome Biol. 3 (12), RESEARCH0084 (2002)		
MEDLINE	22426070		
PUBMED	12537573		
REFERENCE	5 (bases 1 to 288451) Adams,M.D., Celinker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.		
AUTHORS	Direct Submission Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE	6 (bases 1 to 288451) FlyBase Direct Submission Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA		
AUTHORS	7 (bases 1 to 288451) FlyBase Direct Submission Submitted (10-MAR-2004) FlyBase, Harvard University, Biological Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA On Mar 15, 2004 this sequence version replaced gi:21626512.		
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE	1. 288451 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /chromosome="2R" complement (6807, 41460) /locus_tag="CG13499" /map="58B1-58B1" /db_xref="FlyBase:FBgn0034680" complement (FlyBase:FBgn0034680, 7177,39294, 39621,41306, 41460) /locus_tag="CG13499" /product="CG13499-RB" /db_xref="FlyBase:FBgn0034680" complement (join(6807, 7177,39294, 39621,41306, 41460) /locus_tag="CG13499" /product="CG13499-PA" /protein_id="AA046785.2" /db_xref="GI:21626513" /translation="MHQSKSSKSSMMNQYRNQNNKIKQSGSGVTRPTGRNKGAAAGSALNGLMLLLQLVGERLATPSPMOPKRGSGSGGGGGGGGGSSGSSSDVSRDCKNTVDIFEDISPEVSNQVNGVPTCTWVPRITKGA.PRFVRLRRKKRYKYLNLNTHREGTLYQVGVGQPTDLDLPHYSRLTYIG .41460) complement (join(6885, 7164,39288, 39621,41306, 41460) /locus_tag="CG13499" /product="CG13499-RA" /db_xref="FlyBase:FBgn0034680" complement (join(6885, 7164,39288, 39618,40998, 41460) /locus_tag="CG13499" /product="CG13499-RD" /db_xref="FlyBase:FBgn0034680" complement (join(6885, 7164,39288, 39618,41258, 41333) /locus_tag="CG13499" /product="CG13499-RC" /db_xref="FlyBase:FBgn0034680" complement (join(7010, 7164,39288, 39557) /locus_tag="CG13499" /codon_start=1 /product="CG13499-PA" /protein_id="AA068195.1" /db_xref="GI:21626514"		

```
/db_xref="FLYBASE:FBgn0034680"
/translaction="MHOSKSSSEMNQYRNQNNKNQOSGSSGVTRPTMGKNGKAAA
GSRALNCGILMLLLIQLVGERLATPSPSMCPQPKRGREVEAAAEAAVEAVAA
RRTYRRTATRRRTSSRTSARRSATRTVVR"
complement (join(7010..7164,39288..39537))
/locus_tag="CG13499"
/codon_start=1
/product="CG13499-PC"
/protein_id="AA068196.1"
/db_xref="GI:21626515"
/db_xref="FLYBASE:FBgn0034680"
/translaction="MHOSKSSSEMNQYRNQNNKNQOSGSSGVTRPTMGKNGKAAA
GSRALNCGILMLLLIQLVGERLATPSPSMCPQPKRGREVEAAAEAAVEAVAA
RRTYRRTATRRRTSSRTSARRSATRTVVR"
complement (join(7010..7164,39288..39537))
/locus_tag="CG13499"
/codon_start=1
/product="CG13499-PD"
/protein_id="AA068197.1"
/db_xref="GI:21626516"
/db_xref="FLYBASE:FBgn0034680"
/translaction="MHOSKSSSEMNQYRNQNNKNQOSGSSGVTRPTMGKNGKAAA
GSRALNCGILMLLLIQLVGERLATPSPSMCPQPKRGREVEAAAEAAVEAVAA
RRTYRRTATRRRTSSRTSARRSATRTVVR"
<36823..37851
/locus_tag="CG9308"
/map="5881-5881"
/db_xref="FLYBASE:FBgn0034681"
<36823..37851
/locus_tag="CG9308"
/product="CG9308-PA"
/protein_id="AA06786.1"
/db_xref="GI:7291358"
/db_xref="FLYBASE:FBgn0034681"
/translaction="MLSAIVLVLPQAAQGFIVTLDAHETMCFYDHANVSDKYTVSF
EWEGGFKDVGVEIAGDDRLHSHKODMGSFPTMAKRGYOLQCDNMGSMTPTKI
LMQPHVAAIAIEFYMDSSKRVDPVIEDATQSMINOLSAUGAVKRGQEPIMRRYKGH
LEVSDEWELRVLMSTIEGPMMLITITAVLEYIKHFEVRKV"
complement (43569..44784)
/locus_tag="CG10138"
/gene="PpD5"
/map="5881-5881"
/db_xref="FLYBASE:FBgn0005778"
complement (43569..44784)
/gene="PpD5"
/locus_tag="CG10138"

Query Match 25.2%; Score 125.4; DB 3; Length 288451;
Best Local Similarity 65.6%; Pred. No. 2.6e-24;
Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

mRNA
gene
CDS
mRNA
gene
CDS
mRNA
gene
CDS
```

```
Db 227852 GATTAATCTAAATAATCAAACTCCGGCTTTGTAGT 227814
|||||
RESULT 8
AC014140/c 250968 bp DNA linear HTG 16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC014140
VERSION AC014140.1 GI:6437195
KEYWORDS HTG; HTGS PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 250968)
Adams, M. and Venter, J.C.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD, USA
COMMENT This sequence was identified as CDW:10213167 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source 1..250968
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 17.0%; Score 84.8; DB 2; Length 250968;
Best Local Similarity 53.9%; Pred. No. 7.9e-13;
Matches 173; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 173 TTGCTGCCAATTAATTTCTGAGTCGTGATCTTTATACATCACTTAGGTATGCT 232
Db 205056 TTTCTCTCCAGTTCGTTCTGGGGTGGCAACCTTGTAATCTCCCTGGACAGACT 204997
Qy 233 ATTTGACGATACACCAAAAATCTAAAAACACGACATGATATTTCTTATTTTC 292
Db 204996 ACCTCGACGACCAACCAAAACGAAACCACTCGCTCATGCTGCGGTGCCATGGCAC 204937
Qy 293 TTGCTATGCTCGGACGCAACCGGTTATGCTTTGGCTAGCGTGTCTAAAGTTCTACA 352
Db 204936 TCGAGATGATTTGACCAAGTCGTGGGATCTTTTTCGTTCAATCGCTTAACACCTCA 204877
Qy 353 TTTGCCAAGCTTGACGCCCAACAATTGATTAATATGATCTAGATGTTANGACATGGT 412
Db 204876 TCGATCAACCAAGACGCGGTGATCGACAGAAAGCCCCCTGGCTGAGGCTGGT 204817
Qy 413 GGTTAAGTTGGGATGATACATGATCAACTTAATTTTGGACGTTATGATGATGT 472
Db 204816 GGCTCGGCTGGGATGATCGGGACCTCATGATGCTCTTCGCTTATGGAAGTGT 204757
Qy 473 TTGCTAAATTTGCCAAGG 493
Db 204756 TTCCCAAGCAATGCCCAAG 204736

RESULT 9
COS83587 2486 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 11345 from Patent WO0111042.
ACCESSION COS83587
VERSION COS83587.1 GI:41644371
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```


Oy	157	CCTCAAAATTAATACATTGGTCGCCCAATTAATTTCTGGAGTCGGTGAATCTTAATACTAC	216
Db	3221	CCACTTGTGCTCATCTTCTCTCGCAGTTGCTTCGGCGGTGGCAACACTTGTACTAC	3280
Oy	217	ACTTTAGTGTATTCCTATATGACGATTAACACCAAAAATCTAATAACACGACATGATA	276
Db	3281	TCCCTGGGACAGACCTCACTCGACGACACACCAAGAAAGCAACACTCCGCTCATCTCG	3340
Oy	277	AGTTTTCCTATATTTTCTTGTAAGCTGGACCTGGCAACGGTTATAGCTTGGCTAGGTC	336
Db	3341	GCGGTGGCCATGGCACTCAGAGATGATTTGACACAGTCGTGGATTTCTTTTCGGTTTATA	3400
Oy	337	TGTTAAAGTTCTACATTTTGGCCCAACGTTACGCCCAATTTGATTAATATGATCTAGA	396
Db	3401	TCCGTTAACACTTTCATCGATCCAACCAABACGCCCTGATTCGACAGCAAGACCCCCG	3460
Oy	397	TGGTTANGACATGGTGGTTAAGTTGGGATATCTAGATCAACTTAAATATTTTTTGA	456
Db	3461	TGGCTAGAGTCGTGGTGGCTCGGCTGGGTGATCTCGGACACCCATGTCGCTCTTCTCC	3520
Oy	457	ACGTTGATTGGATTGTTTCTCTAAATTCGCGCAAG	493
Db	3521	GGTCTTATCGAGACTGTCTTCCCAAGCATTCATCCCAAG	3557
RESULT 11			
AC009214			
LOCUS	AC009214	154772 bp	DNA linear INV 31-Jul-2004
DEFINITION	Drosophila melanogaster clone BACR05E09, complete sequence.		
ACCESSION	AC009214		
VERSION	AC009214.8	GI:50872371	
KEYWORDS	HTG.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
	Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 154772)		
AUTHORS	Celniker,S.E., Agbayan,A., Arcina,T.T., Baxter,E., Blazej,R.G., Bilenko,C., Champe,M., Chavez,C., Chew,M., Cleislika,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hopkins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomoten,M.A., Mazda,P., Moshtefi,A.R., Moshtefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequelstra,A., Sethi,H., Smit,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA		
REFERENCE	2 (bases 1 to 154772)		
AUTHORS	Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R., Hopkins,R., Stapleton,M., Pacleb,J., Park,S., Svirskas,R., Smith,E., Yu,C. and Rubin,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US		
COMMENT	On Jul 31, 2004 this sequence version replaced gi:15451504. Sequence submitted by: Lawrence Berkeley National Laboratory Berkeley, CA 94720 This sequence submission incorporates changes made during reevaluation of the assembly or fingerprint verification of the clone. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (http://www.fruitfly.org/sequence/) or send email to dbgap@fruitfly.org . location/Qualifiers		
FEATURES	1..154772		
SOURCE	/organism="Drosophila melanogaster" /mol_type="genomic DNA"		

```

/strain="Y; cn bw sp"
/db xref="Caxon:7227"
/chromosome="2L"
/map="33F-33P"
/clone="BAC clone BACR05E09 (D6E5)"
/clone_id="RPT-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcotRI in
pBacE3.6)"

```

Query Match	Best Local Similarity	16.7%	Score 83.2	DB 3	Length 154772	
Matches 178	Conservative	5.8%	0	Mismatches 159	Indels 0	Gaps 0
QY	157	CCTCAAAATPATACTATTGTCGCCCCAATTAATTTCTGAGTCGCTGATCTTATPACTAC	216			
Db	140248	CCACTGTGTGCATCTTTCTCTCGAGTTCTGTGGGGGTGGGAACCTTGATCTAC	140307			
QY	217	ACTTTAGGTGATCCTATATGGACATTAACCCAAAAATCTPAAAACACGACCTGATA	276			
Db	140308	TCCCTGGACAGACCTTACCTCGACACAACAACCAAGAACCAACCTCGCTCATGCTG	140367			
QY	277	AGTTTTCTTATTTTCTTGATAGTGGACCTGCACCGGTTAGCCCTGGCTACAGTC	336			
Db	140368	GCGGTGGCCATGCGACCTAGATGATGTGACCAAGTCGTGGATTCCTTTTCGGTTTCATA	140427			
QY	337	TGTCMAAGTCTACATTTTCGCCACGTTGACGCCAAACATTTGATATATGATCTTAA	396			
Db	140428	TGCGTTAACACTTTCATCGATCCAAACAAGACGCGTTGATGACAGCAAGACCCCCGC	140487			
QY	397	TGTTTANAGCATGTGTGTTAAAGTTGGGTGATCTAGATCAACTTTAATATTTTGGCA	456			
Db	140488	TGGCTAGGTGCTGGTGCTCGGCTGGGGATCTGGGACACCTCAATGTGCTCTTCTCC	140547			
QY	457	ACGTTGATGGATGTTCTTCTPAAATTCGCCAAG	493			
Db	140548	GGTCTTATCGACCTGTTTCCCAAGCAATTGCCCCAAG	140584			

RESULT	12
LOCUS	AC005711/c
DEFINITION	Drosophila melanogaster clone BACR48U09, complete sequence.
ACCESSION	AC005711
VERSION	AC005711.2
KEYWORDS	GI:50872348
SOURCE	HMG.
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha; Epiphytoidae; Drosophilidae; Drosophila. 1 (bases 1 to 165267) Celniker,S.E., George,R.A., Galle,R., Svirskaas,R.R., Hoskins,R.A., Agbayani,A., Archina,T.T., Baker,E., Blazey,R.G., Chavez,C., Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A., Humattci,S.R., Kaira,K., Kearney,L., Kim,S.H., Lee,B., Lomcharin,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitelaw,K.R., Yee,A., Zhang,R., Zieran,L.U. and Kimmel,B.E. Direct Submission Submitted (26-SEP-1998) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA 2 (bases 1 to 165267) Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R., Hoskins,R., Stapleton,W., Pacleb,J., Park,S., Svirskaas,R., Smith,E., Yu,C. and Rubin,G. Direct Submission Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US On Jul 31, 2004 this sequence version replaced gi:4056405. Sequence submitted by:
TITLE	JOURNAL
REFERENCE	AUTHORS
COMMENT	

FEATURES
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This sequence submission incorporates changes made during the reevaluation of the assembly or fingerprint verification of the clone. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdgs@fruitfly.org.
location/Qualifiers

ORIGIN	Query Match	16.7%	Score 83.2	DB 3	Length 165267
Best Local Similarity	52.8%	Pred. No. 2.2e-12			
Matches 178	Conservative 0	Mismatches 159	Indels 0	Gaps 0	
QY	157	CCCTCAATAATTAATCTATTTCCTGTCGCCAATTAATTCCTGAGTCGGTGGATCTTATACTAC	216		
Db	143273	CCACTGTGCTCATCTTTCTCTCGACGTTTCGTTTCGGGCGTGGGCAACACCTTGATCTAC	14322		
QY	217	ACTTTAGGTCTATCCCTATATGACGATACCAACAAAATCTAAACACAGCAGCTGATA	276		
Db	143213	TCCTTGAGACAGACTTACCTCGACGACACACCAAGAAAGCAACATCCGCTCAGTCTG	14311		
QY	277	AGTTTTCCTATTTCCTTCGTATGCTCGACCTGCAACCGTTATGCTTGGCTAGCGTC	336		
Db	143153	CGCGTGCCATGGCAGCTCAGAGATGATGGACCAAGTCGGTGGAATCTTTTCGGTTTCATA	1430		
QY	337	TGCTCAAAAGTTCTACATTCCTGCCAAGCTTGACGCGCAACATTAATTAATTAATCTTACA	366		
Db	143093	TCGCTTAAACACTTCAATTCATCCATCCAAACCAACCGCTGTATGACAGCAAGAACCCCCG	1430		
QY	397	TGGTTTANGACATGGTGGTGAATCACTGGGCGATATCTAGATCAACCTTAATATTTTTCGA	456		
Db	143033	TGGCTAGGTGCTCGGTGGCTCGCTGGGTGATCTGGGACCCCTCATGTGCTCTCTTC	1429		
QY	457	ACGTTGATTGATTGTTTCTCTTAAATTTCTGCCAAGG	493		
Db	142973	GGCTTATGGACGTGTTTCCCAAGCAATTCGCCAAG	142937		
RESULT 13					
AE003636/c					
LOCUS	246230 bp	DNA	linear	INV 15-MAR-2004	
DEFINITION	Drosophila melanogaster chromosome 2L, section 45 of 83 of the				
VERSION	complete sequence				
KEYWORDS	AE003636 AE002690 AE014134				
SOURCE	AE003636.3 GI:22946331				
ORGANISM	Drosophila melanogaster (fruit fly)				
REFERENCE	Drosophila melanogaster				
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 246230) Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabow,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,				

Bailew, R.M., Baev, A., Baxendale, J., Bayraktaroglu, I., Beasley, E.M.
 Beeson, K.Y., Benos, P.V., Berman, B., Bhanderi, D., Bolshakov, S.,
 Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P.,
 Butta, K.C., Chuesm, D.A., Butler, H., Cadieu, E., Center, A.,
 Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B.,
 Davies, P., de Pablos, B., Delcher, A., Deng, Z., May, A.D., Dew, I.,
 Dietz, S.M., Dodson, K., Dou, L.E., Downes, M., Dugan-Rocha, S.,
 Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C.,
 Ferreira, S., Fleischmann, W., Foster, C., Gabrielian, A.E., Garg, N.S.,
 Gelbart, W.M., Glasser, K., Glodok, A., Gong, F., Gottlieb, J.H., Gu, Z.,
 Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J.,
 Hernandez, J.R., Houck, J., Hoehn, D., Housston, K.A., Howland, T.J.,
 Wei, M.H., Ibbagman, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z.,
 Kemison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C.,
 Kravitz, S., Kulp, D., Lai, Z., Lasco, P., Lei, Y., Levitsky, A.A.,
 Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C.,
 McLeod, M.P., McPherson, D., Mekulov, G., Misha, N.V., Mobarry, C.,
 Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,
 Murry, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K.,
 Nuskern, D.R., Pacleby, J.M., Palazolo, M., Pittman, G.S., Pan, S.,
 Pollard, J., Puri, V., Reese, M.G., Rehrer, K., Remington, K.,
 Saunders, R.D., Scheeler, F., Shen, H., Shie, B.C., Siden-Kiamos, I.,
 Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spredling, A.C.,
 Stapleton, M., Strong, R., Sun, E., Svitskas, R., Tector, C., Turner, R.,
 Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A.,
 Weinstein, G.M., Weisenbach, J., Williams, S.M., Woodger, T.,
 Wolley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, Y., Yeh, R.F.,
 Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H.,
 Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,
 Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
 The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)

JOURNAL MEDLINE
 PUBMED 20196006
 TITLE
 2 (bases 1 to 246230)
 REFERENCE
 AUTHORS

Celinker, S.E., Wheeler, D.A., Krommiller, B., Carlson, J.W.,
 Halpern, A., Patel, S., Adams, M., Chame, M., Dugan, S.P., Frise, E.,
 Hodgson, A., George, R.A., Hoskins, R.A., Laverly, T., Murthy, D.M.,
 Nelson, C.R., Pacleby, J.M., Park, S., Pfeiffer, B.D., Richards, S.,
 Sodergren, E.J., Svitskas, R., Taber, P.E., Wan, K., Stapleton, M.,
 Sutton, G.G., Venter, C., Weinstein, G., Scherer, S.E., Myers, E.W.,
 Gibbs, R.A. and Rubin, G.M.
 Finishing a whole-genome shotgun: release 3 of the *Drosophila*
melanogaster euchromatic genome sequence
 Genome Biol. 3 (12), RESEARCH0079 (2002)

JOURNAL MEDLINE
 PUBMED 22426065
 TITLE
 3 (bases 1 to 246230)
 REFERENCE
 AUTHORS

Maier, S., Crosby, M.A., Mungall, C.J., Matthews, B.B., Campbell, K.S.,
 Hradecky, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E.,
 Smith, C.D., Tupy, J.L., Whitfield, E.J., Bayraktaroglu, I.,
 Berman, B.P., Betencourt, B.R., Celinker, S.E., de Grey, A.D.,
 Dyrasdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J.,
 Shu, S.O., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M.,
 Rubin, G.M. and Lewis, S.E.
 Annotation of the *Drosophila melanogaster* euchromatic genome: a
 systematic review
 Genome Biol. 3 (12), RESEARCH0083 (2002)

JOURNAL MEDLINE
 PUBMED 12537572
 TITLE
 4 (bases 1 to 246230)
 REFERENCE
 AUTHORS

Kaminker, J.S., Bergman, C.M., Krommiller, B., Carlson, J.,
 Svitskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E.,
 Rubin, G.M., Ashburner, M. and Celinker, S.E.
 The transposable elements of the *Drosophila melanogaster*
 euchromatin: a genomics perspective
 Genome Biol. 3 (12), RESEARCH0084 (2002)

JOURNAL MEDLINE
 PUBMED 12537573
 TITLE
 5 (bases 1 to 246230)
 REFERENCE
 AUTHORS

Adams, M.D., Celinker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,

[illegible]

Db	949	TATATCGATCCCTTCAAAAAGCCACTGATTACCAAAACGATCCGCGCTGGATGGCGCT	1008
OY	409	TGGTGGTTAAGTTGGTGATCTAGATCAACTTTAATTTTGGCAAGTTGATTGA	468
Db	1009	TGGTGGATAGTTGGATTCTCTCACATTCATCTGACTATCTCAGCAGTGTGCGGC	1068
OY	469	TTGTTTCTTAAATTTGCCCAAGAC	494
Db	1069	ATGTTCCCAAAAGAAATGCCAAGGCG	1094

Search completed: February 26, 2005, 19:44:59
 Job time : 2373 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 17:05:10 ; Search time 511 Seconds
(without alignments)
5769.141 Million cell updates/sec

Title: US-10-621-901-26

Perfect score: 498
Sequence: 1 ggcctcggaacagatgcctc.....aaattctgccaagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing filter 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	9 ACC72953	Acc72953 Cat flea
2	497	99.8	498	12 ADL09595	Adl09595 Cat flea
3	478.6	96.1	498	3 AAC93886	Aac93886 Cat flea
4	478.6	96.1	498	12 ADL09993	Adl09993 Cat flea
5	189.4	38.0	3015	4 ABL05971	AbL05971 Drosophila
6	125.4	25.2	6522	4 ABL05970	AbL05970 Drosophila
7	83.2	16.7	2486	4 ABL09403	AbL09403 Drosophila
8	83.2	16.7	6105	4 ABL09402	AbL09402 Drosophila
9	72.4	14.5	5349	4 ABL06040	AbL06040 Drosophila
10	65.8	13.2	2184	4 ABL06037	AbL06037 Drosophila
11	65.8	13.2	4399	4 ABL06036	AbL06036 Drosophila
12	59	11.8	2552	4 ABL06041	AbL06041 Drosophila
13	41.8	8.4	2172	10 ADC85132	Adc85132 Ion trans
14	41.8	8.4	2172	10 ADC85132	Adc85132 Ion trans
15	41.8	8.4	2175	12 ADM91341	Adm91341 Human tra
16	41.8	8.4	2251	10 ADC85155	Adc85155 Ion trans
17	41.8	8.4	2634	10 ADD37434	Add37434 Human tra
18	41.8	8.4	2634	13 ADR40142	Adr40142 Human OAT
19	41.8	8.4	5054	13 ADR08230	Adr08230 Full leng
20	40.2	8.1	8866	4 AAS45433	Aas45433 Chemical1

21	40.2	8.1	8866	6 ABK28280	Abk28280 DNA trans
22	39.6	8.0	5882	6 ABL32545	AbL32545 Human imm
23	39.4	7.9	110000	3 AAF22305_03	Continuation (4 of
24	38.8	7.8	110000	2 AAT42063_14	Continuation (15 o
25	38.8	7.8	110000	2 AAT42063_15	Continuation (16 o
26	38.6	7.8	1737	10 ADC85178	Adc85178 Ion trans
27	38.6	7.8	2166	10 ADC85157	Adc85157 Ion trans
28	38.6	7.8	2169	10 ADC85181	Adc85181 Ion trans
29	38.4	7.7	15295	4 AAL37039	Aal37039 Human mus
30	38.4	7.7	15295	8 ABX60027	Abx60027 cDNA enco
31	38.4	7.7	15295	12 ADJ30777	Adj30777 Human mus
32	38.2	7.7	2000	8 ADA71938	Ada71938 Rice gene
33	37.6	7.6	3020	10 ADB54000	Adb54000 Human pro
34	37.4	7.5	443	6 ABL94091	AbL94091 Arabidops
35	37.4	7.5	3831	3 AAC48177	Aac48177 Arabidops
36	37.4	7.5	4746	2 AAV57455	Aav57455 Arabidops
37	37.4	7.5	4746	5 AAD03790	Ad03790 Arabidops
38	37.4	7.5	4746	10 ADD44387	Add44387 Arabidops
39	37.4	7.5	4747	2 AAT09019	Aat09019 Arabidops
40	37.4	7.5	4758	12 ADQ13602	Adq13602 Thalecres
41	37.4	7.5	6022	10 ADD44388	Add44388 Arabidops
42	37.4	7.5	6022	12 ADQ13603	Adq13603 Thalecres
43	37.4	7.5	6042	2 AAT09018	Aat09018 Arabidops
44	37.4	7.5	6172	2 AAV57454	Aav57454 Arabidops
45	37.4	7.5	6172	5 AAD03789	Ad03789 Arabidops

ALIGNMENTS

RESULT 1
ACCT2953 standard; DNA; 498 BP.

ACCT2953; 14-JUL-2003 (first entry)

DE Cat flea hindgut and Malpighian tube related gene #11.

XX Antiparasitic; gene therapy; vaccine; cat flea; head; nerve cord;

KW hindgut; malpighian tube; infestation; ss.

XX Ctenocephalides felis.

OS W02003031577-A2.

PD 17-APR-2003.

PF 04-OCT-2002; 2002WO-US031878.

PR 10-OCT-2001; 2001US-0328347P.

PA (HESK-) HESKA CORP.

PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

DR WPI, 2003-393439/37.

PT New nucleic acid molecule, useful for preparing a composition for treating or preventing flea infestation in a mammal.

PS Claim 1; Page 76; 247pp; English.

XX The invention relates to the isolation of nucleic acid sequence from cat fleas that encode head, nerve cord, hindgut and malpighian tube proteins or sequences that hybridize to these. The nucleic acids are useful for preparing a composition for treating or preventing flea infestation in a mammal. The proteins can also be used to raise antibodies for the treatment of flea infestations in mammals. This sequence represents one of the isolated sequences

XX Sequence 498 BP; 148 A; 102 C; 103 G; 144 T; 0 U; 1 Other;

Query Match 99.8%; Score 497; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.7e-139;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCTGGACAAAGATGCCCTTCCAGCTCACCAACCAATATGAGCTGTCTACGATCAAAAT 60
DB 1 GGGCCCTGGACAAAGATGCCCTTCCAGCTCACCAACCAATATGAGCTGTCTACGATCAAAAT 60
QY 61 GCCACCAATGAAGTTCTTAATTAACAGAGAAGAACTTTGTCACACATGATGGCGAA 120
DB 61 GCCACCAATGAAGTTCTTAATTAACAGAGAAGAACTTTGTCACACATGATGGCGAA 120
QY 121 GGAAGTGAATGCGAAGACAGAAAGGAGAGATGCGACCTCAAAATATATATATTTGCTGCC 180
DB 121 GGAAGTGAATGCGAAGACAGAAAGGAGAGATGCGACCTCAAAATATATATATTTGCTGCC 180
QY 181 CAATTAATTTCTGAGATCGGTGATCTTTATATACCTTAAGTATCTATATGAC 240
DB 181 CAATTAATTTCTGAGATCGGTGATCTTTATATACCTTAAGTATCTATATGAC 240
QY 241 GATTAACCAAAAAATCTAAACACAGACCTGATTAAGTTTCTTATTTTCTTCGATG 300
DB 241 GATTAACCAAAAAATCTAAACACAGACCTGATTAAGTTTCTTATTTTCTTCGATG 300
QY 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTAAAGTTCTACATTTCCGCA 360
DB 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTAAAGTTCTACATTTCCGCA 360
QY 361 ACCTTGACCGCAACATTTGATATATATATATATATATATATATATATATATATATAT 420
DB 361 ACCTTGACCGCAACATTTGATATATATATATATATATATATATATATATATATATAT 420
QY 421 TGGGTATCTAGGATCACTTTATATATATTTTGGCAAGTATGATTTGTTCTTAA 480
DB 421 TGGGTATCTAGGATCACTTTATATATATTTTGGCAAGTATGATTTGTTCTTAA 480
QY 481 ATTCTGCAAGAGCTGCA 498
DB 481 ATTCTGCAAGAGCTGCA 498

RESULT 2

ADL09595
ID ADL09595 standard; cDNA; 498 BP.

XX
AC ADL09595;

XX
DT 01-JUL-2004 (first entry)

XX
DE Cat flea hindgut and malpighian tubule (HMT) protein cDNA #12.

XX
KW Flea; head and nerve cord protein; HNC;
KW hindgut and malpighian tubule protein; HMT; flea infestation;
KW anti-arthropod vaccine; chemotherapeutic drug; insecticide; gene; ss;
KW cat flea.

XX
OS Ctenocephalides felis.

XX
PN US2004067516-A1.

XX
PD 08-APR-2004.

XX
PF 16-JUL-2003; 2003US-00621901.

XX
PR 22-JUL-2002; 2002US-0319414P.

XX
PA (BRAN/) BRANDT K S.

XX
PA (GAIN/) GAINES P J.

XX
PA (STIN/) STINCHOMB D T.

XX
PA (WISN/) WISNIEWSKI N.

XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX
DR WPI: 2004-304579/28.
XX
PT Novel flea head and nerve cord protein and flea hindgut and malpighian
PT tubule protein, useful for reducing flea infestations.
PS Claim 5; SEQ ID NO 26; 35pp; English.

CC The invention relates to a flea head and nerve cord (HNC) protein and a
CC flea hindgut and malpighian tubule (HMT) protein. The invention also
CC relates to an isolated nucleic acid molecule expressed by a tissue chosen
CC from a flea HMT tissue and a flea HNC tissue, identified by a method
CC involving constructing a cDNA library enriched for HMT or HNC expressed
CC sequences and identifying a nucleic acid molecule in the library, and an
CC isolated antibody that selectively binds an HNC or HMT protein. The
CC proteins are useful for identifying compounds capable of inhibiting
CC activity of the proteins which involves contacting a protein with a
CC putative inhibitory compound under conditions in which, in the absence of
CC the compound, the protein has activity, and determining if the putative
CC inhibitory compound inhibits its activity. The protein, nucleic acids
CC and antibodies are useful for reducing flea infestations. The proteins
CC and nucleic acids are useful as targets for anti-arthropod vaccines and
CC chemotherapeutic drugs. The proteins are useful for producing a
CC recombinant protein vaccine to protect an animal from flea infestation.
CC The antibodies are useful for passively immunising an animal in order to
CC protect the animal from fleas, as tools to screen expression libraries
CC and/or for recovering desired proteins from a mixture of proteins and
CC other contaminants. The antibodies are also useful for targeting
CC cytotoxic agents to fleas in order to directly kill such fleas. This
CC sequence represents cDNA encoding a cat flea HMT protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPRO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 498 BP; 148 A; 102 C; 103 G; 144 T; 0 U; 1 Other;

Query Match 99.8%; Score 497; DB 12; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.7e-139;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCCTGGACAAAGATGCCCTTCCAGCTCACCAACCAATATGAGCTGTCTACGATCAAAAT 60
DB 1 GGGCCCTGGACAAAGATGCCCTTCCAGCTCACCAACCAATATGAGCTGTCTACGATCAAAAT 60
QY 61 GCCACCAATGAAGTTCTTAATTAACAGAGAAGAACTTTGTCACACATGATGGCGAA 120
DB 61 GCCACCAATGAAGTTCTTAATTAACAGAGAAGAACTTTGTCACACATGATGGCGAA 120
QY 121 GGAAGTGAATGCGAAGACAGAAAGGAGAGATGCGACCTCAAAATATATATATTTGCTGCC 180
DB 121 GGAAGTGAATGCGAAGACAGAAAGGAGAGATGCGACCTCAAAATATATATATTTGCTGCC 180
QY 181 CAATTAATTTCTGAGATCGGTGATCTTTATATACCTTAAGTATCTATATGAC 240
DB 181 CAATTAATTTCTGAGATCGGTGATCTTTATATACCTTAAGTATCTATATGAC 240
QY 241 GATTAACCAAAAAATCTAAACACAGACCTGATTAAGTTTCTTATTTTCTTCGATG 300
DB 241 GATTAACCAAAAAATCTAAACACAGACCTGATTAAGTTTCTTATTTTCTTCGATG 300
QY 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTAAAGTTCTACATTTCCGCA 360
DB 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTAAAGTTCTACATTTCCGCA 360
QY 361 ACCTTGACCGCAACATTTGATATATATATATATATATATATATATATATATATATAT 420
DB 361 ACCTTGACCGCAACATTTGATATATATATATATATATATATATATATATATATATAT 420
QY 421 TGGGTATCTAGGATCACTTTATATATTTTGGCAAGTATGATTTGTTCTTAA 480
DB 421 TGGGTATCTAGGATCACTTTATATATTTTGGCAAGTATGATTTGTTCTTAA 480
QY 481 ATTCTGCAAGAGCTGCA 498

Db 481 ATTCTGCCAAGAGCTGCA 498

RESULT 3
AAC3886
ID AAC3886 standard; cDNA; 498 BP.

XX AAC3886;

XX 19-FEB-2001 (first entry)

DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:381.

XX Cat flea, hindgut and Malpighian tubule nucleic acid; HMT;

KM flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;

XX detection; ss.

XX Ctenocephalides felis.

XX MO200061621-A2.

XX 19-OCT-2000.

XX 07-APR-2000; 2000MO-US009437.

XX 09-APR-1999; 99US-0128704P.

XX (HESK-) HESKA CORP.

XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX WPI; 2000-656323/63.

XX flea Malpighian tubule and head and nerve cord tissue derived nucleic

XX acids useful for the prevention, diagnosis and treatment of flea

XX infestations.

XX Claim 26; Page 372; 964pp; English.

XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic

XX acids which are expressed in hindgut and Malpighian tubule (HMT) tissue

XX or head and nerve cord (HNC) tissue. The invention also relates to the

XX encoded proteins. The invention additionally encompasses expression

XX constructs, recombinant viruses and recombinant cells comprising the

XX nucleic acids of the invention, recombinant production of the proteins,

XX antibodies against the proteins, a method of identifying inhibitors of

XX administration to an animal. The nucleic acids, and the proteins they

XX encode may be used in the prevention, treatment and diagnosis of diseases

XX associated with flea infestations. For example, the nucleic acids may be

XX used to produce an HMT or HNC protein according to standard recombinant

XX DNA methodology by inserting the nucleic acids into a host cell and

XX culturing the cell to express the protein. The HMT and HNC nucleic acids

XX may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect

XX and quantitate the presence of cat flea or other homologous nucleic acid

XX sequences in samples. They may also be used to study the expression and

XX function of the proteins and their role in metabolism. The HMT and HNC

XX proteins may be used as antigens in the production of specific

XX antibodies, and in assays to identify modulators (agonists and

XX antagonists) of HMT and/or HNC protein expression and activity. The anti-

XX HMT/HNC protein antibodies and antagonists may also be used to

XX downregulate protein expression and activity. The antibodies may also be

XX used as diagnostic agents for detecting the presence of flea polypeptides

XX in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The

XX present sequence represents a cat flea HMT cDNA of the invention

XX

XX

XX

XX

QY 1 GGCCCTGGAACAGATGCCCTTTCGACTCACCACCGAATATGAGCTGTCTACGATCAAAAT 60

Db 1 GGCCCTGGAACAGATGCCCTTTCGACTCACCACCGAATATGAGCTGTCTACGATCAAAAT 60

QY 61 GCCACCATTTGAAGTTCTTAATTAACAGAGAGGAAAACTTTGGCCACACTGATGGGAA 120

Db 61 GCCACCATTTGAAGTTCTTAATTAACAGAGAGGAAAACTTTGGCCACACTGATGGGAA 120

QY 121 GGAATTGAATGCAAGAAAGAGGAGCATGCACCTCAATTAATCTATTGTGCGCC 180

Db 121 GGAATTGAATGCAAGAAAGAGGAGCATGCACCTCAATTAATCTATTGTGCGCC 180

QY 181 CAATTAATTTCTGAGTGGTGTGATCTTTATATACACTTATAGTGTATCTATATGAC 240

Db 181 CAATTAATTTCTGAGTGGTGTGATCTTTATATACACTTATAGTGTATCTATATGAC 240

QY 241 GATTAACCAAAAAATCTAAACACACGACTGATTAAGTTTCTTATTTCTTCGATG 300

Db 241 GATTAACCAAAAAATCTAAACACACGACTGATTAAGTTTCTTATTTCTTCGATG 300

QY 301 CTGGACCTGCAACCGGTTATGCTTGGCTAGGAGTGTCTAAAGTTCTACATTTGCGCA 360

Db 301 CTGGACCTGCAACCGGTTATGCTTGGCTAGGAGTGTCTAAAGTTCTACATTTGCGCA 360

QY 361 ACGTTGACGCCAACAAATTGATTAATGATCTAGATGTTAGAGCATGTGGTTAAGT 420

Db 361 ACGTTGACGCCAACAAATTGATTAATGATCTAGATGTTAGAGCATGTGGTTAAGT 420

QY 421 TGGGTGATTAAGATCACTTAATATTTTGGCAAGTTGATGTTGTTCTTAA 480

Db 421 TGGGTGATTAAGATCACTTAATATTTTGGCAAGTTGATGTTGTTCTTAA 480

QY 481 ATTCTGCCAAGAGCTGCA 498

Db 481 A-TCTGCCAAGAGCTGCA 497

RESULT 4

ID ADL09993

XX ADL09993 standard; cDNA; 498 BP.

XX ADL09993;

XX 01-JUL-2004 (first entry)

XX Cat flea hindgut and malpighian tubule (HMT) protein cDNA #410.

XX flea; head and nerve cord protein; HNC;

XX hindgut and malpighian tubule protein; HMT; flea infestation;

XX anti-arthropod vaccine; chemotherapeutic drug; insecticide; gene; ss;

XX cat flea.

XX Ctenocephalides felis.

XX US2004067516-A1.

XX 08-APR-2004.

XX 16-JUL-2003; 2003US-00621901.

XX 22-JUL-2002; 2002US-0319414P.

XX (BRAN/) BRANDT K S.

XX (GAIN/) GAINES P J.

XX (STIN/) STINCHCOMB D T.

XX (WISN/) WISNIEWSKI N.

XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX WPI; 2004-304579/28.

XX Novel flea head and nerve cord protein and flea hindgut and malpighian

XX tubule protein, useful for reducing flea infestations.

XX

XX

XX

XX

XX Claim 5; SEQ ID NO 426; 35pp; English.

CC The invention relates to a flea head and nerve cord (HNC) protein and a
CC flea hindgut and malpighian tubule (HMT) protein. The invention also
CC relates to an isolated nucleic acid molecule expressed by a tissue chosen
CC from a flea HMT tissue and a flea HNC tissue, identified by a method
CC involving constructing a cDNA library enriched for HMT or HNC expressed
CC sequences and identifying a nucleic acid molecule in the library, and an
CC isolated antibody that selectively binds an HNC or HMT protein. The
CC proteins are useful for identifying compounds capable of inhibiting
CC activity of the proteins which involves contacting a protein with a
CC putative inhibitory compound under conditions in which, in the absence of
CC the compound, the protein has activity, and determining if the putative
CC inhibitory compound inhibits its activity. The proteins, nucleic acids
CC and antibodies are useful for reducing flea infestations. The proteins
CC and nucleic acids are useful as targets for anti-arthropod vaccines and
CC chemotherapeutic drugs. The proteins are useful for producing a
CC recombinant protein vaccine to protect an animal from flea infestation.
CC The antibodies are useful for passively immunizing an animal in order to
CC protect the animal from fleas, as tools to screen expression libraries
CC and/or for recovering desired proteins from a mixture of proteins and
CC other contaminants. The antibodies are also useful for targeting
CC cytotoxic agents to fleas in order to directly kill such fleas. This
CC sequence represents cDNA encoding a cat flea HMT protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

SO Sequence 498 BP; 147 A; 101 C; 105 G; 145 T; 0 U; 0 Other;

Query Match 96.1%; Score 478.6; DB 12; Length 498;
Best Local Similarity 98.8%; Pred. No. 1.3e-133;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATGAGTGTCTAGATCAAAAT 60
DB 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATGAGTGTCTAGATCAAAAT 60
QY 61 GCCACCATTTGAAGTTCTTAATAAACAAGAGAGAACTTTGTCACACATGAGCGAA 120
DB 61 GCCACCATTTGAAGTTCTTAATAAACAAGAGAGAACTTTGTCACACATGAGCGAA 120
QY 121 GGAAGTGAATGCGAAACAGAAAGAGAGATGCGACCTCAATAATATCTATTTGCTGCC 180
DB 121 GGAAGTGAATGCGAAACAGAAAGAGAGATGCGACCTCAATAATATCTATTTGCTGCC 180
QY 181 CAATTATTTCTGAGTGGTGGATCTTTATCTACGCTTAGTGATCTATATGAGAC 240
DB 181 CAATTATTTCTGAGTGGTGGATCTTTATCTACGCTTAGTGATCTATATGAGAC 240
QY 241 GATTAACACCAAAAATCTAATAACACGACATGATGTTTCTTATTTCTTCGTATG 300
DB 241 GATTAACACCAAAAATCTAATAACACGACATGATGTTTCTTATTTCTTCGTATG 300
QY 301 CTGGAACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTTAAGTTCTACATTTGCCA 360
DB 301 CTGGAACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTTAAGTTCTACATTTGCCA 360
QY 361 ACGTGAACGCAACAATTTGAATATGATCTAGAGTGTAGAGATGATGATGATGATGAT 420
DB 361 ACGTGAACGCAACAATTTGAATATGATCTAGAGTGTAGAGATGATGATGATGATGAT 420
QY 421 TGGGATGATCTAGAGTCACTTTAATATTTTTCGACAGTTGATGATGATGATGATGAT 480
DB 421 TGGGATGATCTAGAGTCACTTTAATATTTTTCGACAGTTGATGATGATGATGATGAT 480
QY 481 ATTTCGCAAGAGCTGCA 498
DB 481 ATTTCGCAAGAGCTGCA 497

RESULT 5

ABL05971
ID ABL05971 standard; cDNA; 3015 BP.

XX ABL05971;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12395.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001MO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656860/75.

XX P-P-SDB; ABB61868.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

PS Claim 1; SEQ ID NO 12395; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL0175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences

SO Sequence 3015 BP; 754 A; 741 C; 799 G; 721 T; 0 U; 0 Other;

Query Match 38.0%; Score 189.4; DB 4; Length 3015;
Best Local Similarity 61.3%; Pred. No. 3.3e-46;
Matches 304; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATGAGTGTCTAGATCAAAAT 60
DB 1127 GGAACCGGCGAGAGTCCACTGCTGACCTGGAATGCGCGGAGTCCCGATGAAAAT 1186
QY 61 GCCACCATTTGAAGTTCTTAATAAACAAGAGAGAACTTTGTCACACATGAGCGAA 120
DB 1187 GCCACCATTTGAAGTTCTTAATAAACAAGAGAGAACTTTGTCACACATGAGCGAA 1246
QY 121 GGAAGTGAATGCGAAACAGAAAGAGAGATGCGACCTCAATAATATCTATTTGCTGCC 180
DB 1247 GGAAGTGAATGCGAAACAGAAAGAGAGATGCGACCTCAATAATATCTATTTGCTGCC 1306
QY 181 CAATTATTTCTGAGTGGTGGATCTTTATCTACGCTTAGTGATCTATATGAGAC 240
DB 1307 CAGTTATATCTGAGATTTGAGATGCTATATCTACGCTTAGTGATGATGATGATGATGAT 1366
QY 241 GATTAACACCAAAAATCTAATAACACGACATGATGTTTCTTATTTCTTCGTATG 300
DB 1367 GATTAACACCAAAAATCTAATAACACGACATGATGTTTCTTATTTCTTCGTATG 1426

[illegible]

Db	3510	GGACCGGGCGAGGATGCACTTGCTCTGACCTCGGAGTTGGCGGGATGCCCCGATGAAAT	3451
Qy	61	GCCACCATTTGAAGTCTTAATTAACAGAAAGAGAAATTTGTGCCACATGATGGGAA	120
Db	3450	GCCCAATAGAGGCTATCGAGAGACGAGATCCAAAGACCTTATCGCGTTGAATGGAGGA	3391
Qy	121	GGAGTTGAATCCGAAACAGAAAGAGGAGATCGCACCTCAATTAATACTATTTGCTGCC	180
Db	3390	GGAGCGAGTGCAGAGTCGATCGTAGAAGAAACTTTGGCCCCCAACTTTTGCTCTTGATGCG	3311
Qy	181	CAATTAAATTTCTGGAGTCGGTGATCTTTTAATCACTTTAAGTGATCTCTAATGGAC	240
Db	3330	CAGTTCAATCTGGAAATGGAGATCGCTAATCAACACGTTGGGAGATGCTCATATGAC	3211
Qy	241	GATAACACCAAAATCTTAAACACGACCATGTAAAGT	279
Db	3210	GATTAATCAAAAAATCCAAAATCTCGGCTTTGTGGAGT	3232

ID	ABL09403	standard; cDNA; 2486 BP.
AC	ABL09403;	
DT	26-MAR-2002	(first entry)
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 22691.	
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
OS	Drosophila melanogaster.	
PN	WO200171042-A2.	
PD	27-SEP-2001.	
XX	23-MAR-2001; 2001WO-US009221.	
PF	23-MAR-2000; 2000US-0191637P.	
PR	11-JUL-2000; 2000US-00614150.	
XX	(PEXE) PE CORP NY.	
PA	Venter JC, Adams M, Li PWD, Myers EW;	
PI	WPI; 2001-656660/75.	
DR	P-PsDB; ABB65300.	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.	
PT	Claim 1; SEQ ID NO 22691; 21pp + Sequence Listing; English.	
PS	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLU6176-ABU30511), expressed DNA sequences (ABU01840-ABU6175) and the encoded proteins (ABB57737-ABR72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from Wipo at ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 2486 BP; 565 A; 717 C; 621 G; 583 T; 0 U; 0 Other;	
SQ	Query Match	16.7%; Score 83.2; DB 4; Length 2486;
	Best Local Similarity	54.0%; Pred. No. 2,9e-14;
	Matches 169; Conservative	0; Mismatches 144; Indels 0; Gaps 0;

```

RESULT 8
ABLO9402
ID      ABL09402 standard; cDNA; 6105 BP
XX
AC      ABL09402;

```

PS Claim 1, SEQ ID NO 22686; 21np + Sequence Listing; English.

QY	157	UCTCAAAATAACTACTTTGGCTGCCCAATTAATTTCTGGATCGGTGGATCTTTATATAC	216
Db	3221	CCACTTGGCTCACTTCTTCTCTGCGAGTTGGTTCCTGGGCGTGGGCAACACTGTGTACTAC	3280
QY	217	ACCTTAGGTGATCTCATATGAGACGATACACCCAAAATCTTAAACACCGACGACTGATA	276
Db	3281	TCCTCGGACAGACCTTACCTCGAGGACAAACCAAGAAAGCGAACTTCGGCTCATCTCG	3340
QY	277	AGTTTTCTTATTTTCTTCGTATGCTCGGACCTGGCAACCGATTATGCTTGGCTAGCGTC	336
Db	3341	GGGGTGGCCATGGGACCTCAGGATGATTGGACGAGTCGTGGGATTCTTTTCGGTTTCATA	3400
QY	337	TGCTTAAAGTTCTACATTTGCGCAACGTTGACGCCAACAATTGATTAATTAATGATCTAGA	396
Db	3401	TCGCTTAAACACTTTCATTCGATTCGAACCAAGACCGCGTTGATCGACACAGAGACCCCGC	3460
QY	397	TGCTTANAGCATGCTGGTTAAAGTTGGGTGATCTAGAGATCAACTTAAATATTTTTTGA	456
Db	3461	TGGCTAGGTGCTGTGGTCTGCGCTGGGTGATCTCTGGGCAACCTCATGTGCTCTCTCC	3520
QY	457	ACGTTGATGGATTTGTTTCTTAAATATCTGCCAAG	493
Db	3521	GGTCTTATCGGACCTGTTCCCAAGCAATTTGCCCAAG	3557

RESULT 9
 ABL06040/C
 ID ABL06040 standard; cDNA; 5349 BP.
 XX
 AC ABL06040;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12602.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEXE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR P-PsDB; ABB61937.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 12602; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB557737-
CC AB572072). The sequence data for this patent did not form part of
CC printed specification, but was obtained in electronic format direct
CC from WIPO at http://wipo.int/pub/published_pct_sequences

SQ Sequence 5349 BP; 1439 A; 1283 C; 1159 G; 1468 T; 0 U; 0 Other;

Query Match	14.5%	Score	72.4	DB	4	Length	5349
Best Local Similarity	58.0%	Pred. No.	7e-11				
Matches	127	Conservative	0	Mismatches	92	Indels	0
						Gaps	0

Oy 276 AAGTTTTCTATTTCTCGTAGCTCGAACCGCAACGGTATGCGTTGGCTAGCGT 335
 Db 3451 AGGCATGCTACCTTCTAAGAGATCTGGTCAAGCCATTGGTATCTCCCTAGCTGCT 3392

OY	336	CTGTCTAAAGTTCTACATTTGCGCAACGTGACGGCAACAATTATTAATATGATCTCTTG	395
Db	3391	CTTGCTCGCGCCTATAACATTGAGCCACCTTGAGGCACGTATGGCAGGAAGAATCCACG	3332

Oy 386 ATGCTTANAGCATGTGGTTACGTGGTGATPACTAGCATCAACTTAAATTTTTTC 455
3331 CTGGTGGCGCCCTGGTGGCGGTGGCTGTCTCTGGCGCAATTACCCTGGTCTCCGC 3272

QY 456 AACGTTGATGATGTTTCTAAATTCGCCAGAGC 494
| | | | | | | | | | | | | | | | | |
Db 3271 CATTCGTGTTTAATGTTCCCAAGCAGCTGCCAGTGC 3233

RESULT 10
ABL06037
ID ABL06037 standard; cDNA; 2184 BP.

AC	ABL06037;
XX	
DT	26-MAR-2002 (first entry)

DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 12593
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;

XX *Drosophila melanogaster*.
XX

EN	NO2001/1012-22
XX	
PD	27-SEP-2001.
XX	

PF	23-MAR-2001; 2001WC-US009231
XX	
PR	23-MAR-2000; 2000US-0191637P
PR	11-JUL-2000; 2000US-00614150

XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC Adams M

XX WPI; 2001-656860/75.
DR P-PSDB; ABB61934.
DR
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.

PS Claim 1; SEQ ID NO 12593; 21pp + Sequence Listing; English
XX
CC The invention relates to an isolated nucleic acid detection
method of detecting a target nucleic acid sequence.

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC

CC AB372072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2184 BP; 516 A; 511 C; 557 G; 600 T; 0 U; 0 Other;

Query Match	13.2%	Score 65.8	DB 4	Length 2184
Best Local Similarity	57.3%	Pred. No. 4.8e-09		
Matches 118; Conservative	0	Mismatches 88	Indels 0	Gaps 0

Qy 289 TTCTTCGATGCTGGACTGCAACCGGTATACCTTGGCTAGCGTCGTCTAAGTTC 348

Db 889 TTCTCGGATGCTGGGCTCTCTATAGGCTTCTTCATGATGATACCTTGCTCCGGCTG 948

QY 349 TACATTTGGCCACGTTGACGCCAACAAATTGATATATATGATCCTAGATGCTTANAGCA 408

Db 949 TATATGATCCTTCAAAGAAGCCACTGATTACCAACAACGATCCGGCTGGATGGCGCT 1008

QY 409 TGGTGTTAAGTTGGGTGATCTAGAGTCAACTTTAATTTTTTGGCAAGTGCATTGGA 468
|||||
Db 1009 TGGTGAGTAGGTTGGATTCTCTTCACATTCATCTGACTATCTTAGCAAGTGTGTGGC 1068

```

QY      469 TTGTTTCTAAATTCGCGAAGAGC 494
          |||||
Db      1069 ATGTTCCCAAGAAATGCCAAGGGC 1094

```

RESULT 11
ABL06036/c
ID ABL06036 standard; cDNA; 4399 BP.

AC	ABL06036;
XX	
DT	26-MAR-2002 (first entry)

DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 12590.
XX
KW *Drosophila*; developmental biology; cell signalling; insecticide;

XX
OS
XX

Drosophila melanogaster.

XX 27-SEP-2001.
PD
XX

FF	23-MAR-2001; 2001MO-030052231
XX	
PR	23-MAR-2000; 2000US-0191637P.
PR	11-JUL-2000; 2000US-00614150.

AA
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M. J: PMD, Myers EW:

XX WPI; 2001-656860/75
DR P-PSDB; ABB61933.
DR
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
YX

PS Claim 1; SEQ ID NO 12590; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection
CC method of detecting 1000 or more genes from *Neisseria*. The

CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB1016176-AB130511), expressed DNA
CC sequences (AB101840-AB161615) and the encoded proteins (A855737-
CC AB872012). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly

CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 4399 BP; 1243 A; 1082 C; 953 G; 1121 T; 0 U; 0 Other;
 SQ Query Match 13.2%; Score 65.8; DB 4; Length 4399;
 Best Local Similarity 57.3%; Pred. No. 6.3e-09;
 Matches 118; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
 QY 289 TTTCCTGATGCTCGGACCTGCAACCGTTATGCTTGGCTAGCGCTGTCTAAAGTTC 348
 |||||
 DB 2397 TTTCGCGATGCTGGGCTCTGCTATAGGCTTCCATGATGATCTTGTCCTCCGCTG 2338
 |||||
 QY 349 TACATTTCCGCAACGTTGACGCCAATGATTAATATATCTAGATGTTANAGCA 408
 |||||
 DB 2337 TATATGATTCCTTCAAAAAGCCATGATACCAAAACATCCGCGCTGATGGCGCT 2278
 |||||
 QY 409 TTGTTGATGATGTTGGGATGACTAGATCACTTTATATTTTGGCAAGTTGATGA 468
 |||||
 DB 2277 TGGTGAATAGTTGATTTCTCTCACTTCACTTCACTTCAAGCATGTTGCTGGC 2218
 |||||
 QY 469 TTGTTCTTAAATTCGCCAAGAC 494
 |||||
 DB 2217 ATGTTCCCAAGAAATGCCAAGGC 2192
 |||||
 RESULT 12
 ABL06041
 ID ABL06041 standard; cDNA; 2552 BP.
 XX
 AC ABL06041;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12605.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW,
 XX
 DR WPI; 2001-656860/75.
 XX
 DR P-PADB; ABB61938.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 12605; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB10511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 2552 BP; 663 A; 625 C; 678 G; 586 T; 0 U; 0 Other;
 Query Match 11.8%; Score 59; DB 4; Length 2552;
 Best Local Similarity 67.5%; Pred. No. 5.7e-07;
 Matches 83; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 157 CCTCAATATATCTATTTTGTGCTCCCAATTAATTTTCGAGTGGTGAATCTTATCTAC 216
 |||||
 DB 609 CCCCAAGTGTTCCTGTTCTGCTCTCACTAATCTCGGAGTGGTCAAGCTCTGTTCTAT 668
 |||||
 QY 217 ACTTAGGTGATACCTATATGACGATTAACCAAAAATCTTAAACACAGCACTGATA 276
 |||||
 DB 669 ACCCTGGCATGACCTTACATGACGACATACCAAGCAAGCAAGCACTGCAATGCTG 728
 |||||
 QY 277 AGT 279
 |||||
 DB 729 ACT 731
 |||||
 RESULT 13
 ADC85132
 ID ADC85132 standard; DNA; 2172 BP.
 XX
 AC ADC85132;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Ion transporter protein coding sequence, SEQ ID 2.
 XX
 KW Ion transporter protein; kidney disease; liver disease; pancreas disease;
 KW immunological disease; thymus failure; reproductive disease;
 KW digestive disease; spleen disease; cancer; respiratory disease; myelitis;
 KW diabetes; hypertension; reperfusion injury; ischemia; retinitis;
 KW central nervous disease; skin disease;
 KM thyroid hormone-associated disease; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2172
 FT /*tag= a
 FT /product= "Ion transporter protein"
 XX
 PN WO2003074702-A1.
 XX
 PD 12-SEP-2003.
 XX
 PF 05-MAR-2003; 2003WO-JP002564.
 XX
 PR 06-MAR-2002; 2002JP-00061133.
 XX
 PR 01-APR-2002; 2002JP-00098852.
 XX
 PR 25-JUN-2002; 2002JP-00184883.
 XX
 PA (TAKA) TAKEDA CHEM IND LTD.
 XX
 PI Nakaniishi A, Hikichi Y, Uno Y;
 XX
 DR WPI; 2003-722073/68.
 XX
 DR P-PADB; ADC85131.
 XX
 PT Novel marker proteins useful for the treatment and prevention of kidney
 PT disorders.
 XX
 PS Claim 10; SEQ ID NO 2; 186pp; Japanese.
 XX
 CC The present invention relates to novel ion transporter proteins
 CC (ADC85131, ADC85156, ADC85162 and ADC85184) and their coding sequences
 CC (ADC85132, ADC85157, ADC85163 and ADC85185). The proteins are useful as
 CC diagnostic markers for kidney diseases, liver diseases, pancreas
 CC diseases, immunological diseases associating thymus failures,
 CC reproductive diseases, digestive diseases, spleen diseases, cancer,
 CC respiratory diseases, myelitis, diabetes, hypertension, reperfusion
 CC injury following ischemia, retinitis, central nervous diseases, skin

CC diseases and thyroid hormone-associated diseases, and also for the
CC treatment and prevention of these diseases.

XX Sequence 2172 BP; 581 A; 451 C; 451 G; 689 T; 0 U; 0 Other;

Query Match 8.4%; Score 41.8; DB 10; Length 2172;
Best Local Similarity 47.8%; Pred. No. 0.08;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

```
QY 181 CAATTAATTTCTGGAGTCGGTGAATCTTATACACTTATAGTATCTATATGAC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 CAACATATGCTGGGGGAGAGGAACTCTCTTATCTGGGAACGCTTCTTGAT 765
QY 241 GATTAACACCAAAAATATCTAAACACAGCACTGATTAATTTTCTTCTGATG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 GATTCTGTGCCACACACAAAGTCTTCTCTATATAGAACGGTATGATGCAATC 825
QY 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTAAAGTTCAATTCGCCA 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 TTAGGCCCTGCTATGCTATGATGTATGGAGACAACTGATACCATATACATGATGT 885
QY 361 ACGTTAGCCCAAAATTTGA-----TAATATGATCTTATAGTATGATGATGG 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 GCTATGGAGAAAGCACTGATGCTAGAGATGATCCGATGTTGGAGCTTGGTGG 945
QY 415 TTAAGTTGGGTGATACAGATCACTTAATATTTTGGCAAGTTGATGATGTTT 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 ATTGGGTTTCTTCTATCATGATCTTTCGTTCTTAAATATACCTTTTCTGCTTT 1005
QY 475 CTTAAATTTCTGCCAAGAGCTGCA 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 CCAAAACATTTACCAAGTACAGCA 1029
```

RESULT 14
ADD37436
ID ADD37436 standard; cDNA; 2172 BP.

XX AC ADD37436;

XX DT 15-JAN-2004 (first entry)

XX DE Human transporter OAT-5 cDNA #2.

XX KM Human; 861 gene; transporter; cytosolic; anorectic; antidiabetic;

XX KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;

XX OS Obesity; epilepsy; diabetes.

XX PN Homo sapiens.

XX US2003143675-A1.

XX 31-JUL-2003.

XX PF 22-MAY-2002; 2002US-00154419.

XX PR 12-MAY-2000; 2000US-0204211P.

XX PR 29-JUN-2000; 2000US-0215376P.

XX PR 31-JUL-2000; 2000US-0221769P.

XX PR 19-SEP-2000; 2000US-0233790P.

XX PR 25-SEP-2000; 2000US-0235107P.

XX PR 05-OCT-2000; 2000US-0238336P.

XX PR 14-NOV-2000; 2000US-0248354P.

XX PR 15-NOV-2000; 2000US-0248787P.

XX PR 15-DEC-2000; 2000US-0256240P.

XX PR 18-DEC-2000; 2000US-0256588P.

XX PR 21-DEC-2000; 2000US-0258028P.

XX PR 22-JAN-2001; 2001US-0263169P.

XX PR 14-MAY-2001; 2001US-00858194.

XX PR 29-JUN-2001; 2001US-0085811.

XX PR 31-JUL-2001; 2001US-00919781.

XX PR 19-SEP-2001; 2001US-00957664.

XX PR 25-SEP-2001; 2001US-00964295.

PR 05-OCT-2001; 2001US-00972724.

PR 14-NOV-2001; 2001US-00002769.

PR 17-DEC-2001; 2001US-00024623.

PR 22-JAN-2002; 2002US-00055025.

XX (MILL-) MILLENNIUM PHARM INC.

XX Curtis RAJ, Gluckemann MA, Meyers RE;

XX WPI; 2003-851783/79.

XX P-PSDB; ADD37435.

XX New isolated nucleic acid, useful for preparing a composition for

XX treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy

XX or diabetes.

XX Claim 1; SEQ ID NO 9; 663bp; English.

XX The invention relates to an isolated nucleic acid comprising a cDNA

XX encoding a human transporter protein, or its complement, a sequence that

XX is 60 % identical to the cDNA, a fragment comprising at least 30

XX nucleotides of the cDNA, or a sequence encoding a fragment of the

XX polypeptide comprising at least 10 contiguous amino acid residues of the

XX cDNA. Also included are a vector comprising the novel nucleic acid

XX molecule, producing the polypeptide, the isolated transporter

XX polypeptide, an isolated antibody that specifically binds to the

XX polypeptide, detecting the presence of the polypeptide or nucleic acid in

XX a sample, a kit, identifying a compound that binds to, or that modulates

XX the activity of, the polypeptide, and modulating the activity of the

XX polypeptide. The nucleic acid is useful for preparing a composition for

XX treating PGC-1 (not defined) associated disorders e.g. liver tumors,

XX obesity, epilepsy or diabetes. The present sequence encodes a novel human

XX transporter protein.

XX Sequence 2172 BP; 580 A; 450 C; 453 G; 689 T; 0 U; 0 Other;

XX Query Match 8.4%; Score 41.8; DB 10; Length 2172;

XX Best Local Similarity 47.8%; Pred. No. 0.08;

XX Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

QY 181 CAATTAATTTCTGGAGTCGGTGAATCTTATACACTTATAGTATCTATATGAC 240

Db 706 CAACATATGCTGGGGGAGAGGAACTCTCTTATCTGGGAACGCTTCTTGAT 765

QY 241 GATTAACACCAAAAATATCTAAACACAGCACTGATTAATTTTCTTCTGATG 300

Db 766 GATTCTGTGCCACACACAAAGTCTTCTCTATATAGAACGGTATGATGCAATC 825

QY 301 CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTAAAGTTCTACATTCGCCA 360

Db 826 TTAGGCCCTGCTATGCTATGATGTATGGAGACAACTGATACCATATACATGATGT 885

QY 361 ACGTTAGCCCAAAATTTGA-----TAATATGATCTTATAGTATGATGATGG 414

Db 886 GCTATGGAGAAAGCACTGATGCTAGAGATGATCCGATGTTGGAGCTTGGTGG 945

QY 415 TTAAGTTGGGTGATACAGATCACTTAATATTTTGGCAAGTTGATGATGTTT 474

Db 946 ATTGGGTTTCTTCTATCATGATCTTTCGTTGCTTAAATATACCTTTTCTGCTTT 1005

QY 475 CTTAAATTTCTGCCAAGAGCTGCA 498

Db 1006 CCAAAACATTTACCAAGTACAGCA 1029

RESULT 15

ADW91341
ID ADW91341 standard; DNA; 2175 BP.

XX AC ADW91341;

XX DT 03-JUN-2004 (first entry)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 18:50:01 ; Search time 147 Seconds

(without alignments)
5543.305 Million cell updates/sec

Title: US-10-621-901-26

Perfect score: 498

Sequence: 1 ggcctcggaagagtcgcct.....aaattctgcaagagctgca 498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2_6/ptodaca/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodaca/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodaca/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodaca/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodaca/1/ina/PTCUS COMB.seq:*

6: /cgn2_6/ptodaca/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	38.8	7.8	1830121	4 US-09-557-884-1	Sequence 1, Appl1
C 2	38.8	7.8	1830121	4 US-09-643-990A-1	Sequence 1, Appl1
C 3	38.4	7.7	601	4 US-09-949-016-126480	Sequence 126480, A
C 4	38.4	7.7	69642	4 US-09-949-016-15339	Sequence 15339, A
C 5	37.6	7.6	2721	4 US-09-949-016-713	Sequence 713, App
C 6	37.6	7.6	2721	4 US-09-949-016-2163	Sequence 2163, App
C 7	37.6	7.6	2721	4 US-09-949-016-2164	Sequence 2164, App
C 8	37.6	7.6	2741	4 US-09-949-016-1119	Sequence 1119, App
C 9	37.4	7.5	4746	2 US-08-819-288-2	Sequence 2, Appl1
C 10	37.4	7.5	4746	2 US-08-400-348-2	Sequence 2, Appl1
C 11	37.4	7.5	4747	1 US-08-261-822A-2	Sequence 2, Appl1
C 12	37.4	7.5	4747	5 PCT-US95-07744A-2	Sequence 2, Appl1
C 13	37.4	7.5	6042	1 US-08-261-822A-1	Sequence 1, Appl1
C 14	37.4	7.5	6042	5 PCT-US95-07744A-1	Sequence 1, Appl1
C 15	37.4	7.5	6172	2 US-08-819-288-1	Sequence 1, Appl1
C 16	37.4	7.5	6172	3 US-08-400-348-1	Sequence 1, Appl1
C 17	35.4	7.1	269223	4 US-09-556-002-41	Sequence 41, Appl1
C 18	34.6	6.9	168575	3 US-09-426-280-1	Sequence 1, Appl1
C 19	34.2	6.9	601	4 US-09-949-016-195940	Sequence 195940, A
C 20	34.2	6.9	601	4 US-09-949-016-195941	Sequence 195941, A
C 21	34.2	6.9	3692	4 US-09-575-081B-7	Sequence 7, Appl1
C 22	34.2	6.9	3725	4 US-09-799-451-260	Sequence 260, App
C 23	34.2	6.9	231672	4 US-09-949-016-17296	Sequence 17296, App
C 24	34.2	6.9	251682	4 US-09-949-016-11973	Sequence 11973, A
C 25	34.2	6.9	325791	4 US-09-768-185A-1	Sequence 1, Appl1
C 26	34	6.8	150780	4 US-09-949-016-14711	Sequence 14711, A
C 27	33.8	6.8	928	4 US-09-270-767-1100	Sequence 1100, App

C 28	33.8	6.8	928	4 US-09-270-767-16382	Sequence 16382, A
C 29	32.8	6.6	601	4 US-09-949-016-134819	Sequence 134819, App
C 30	32.8	6.6	832	4 US-09-621-976-2813	Sequence 2813, App
C 31	32.8	6.6	903	4 US-09-328-352-3071	Sequence 3071, App
C 32	32.8	6.6	3183	4 US-09-248-796A-3959	Sequence 3959, App
C 33	32.8	6.6	76610	4 US-09-949-016-15521	Sequence 15521, App
C 34	32.8	6.6	125192	4 US-09-949-016-14120	Sequence 14120, A
C 35	32.8	6.6	421491	4 US-09-949-016-12805	Sequence 12805, A
C 36	32.8	6.6	421494	4 US-09-949-016-14060	Sequence 14060, A
C 37	32.8	6.6	678533	4 US-09-949-016-14577	Sequence 14577, A
C 38	32.8	6.6	678533	4 US-09-949-016-14578	Sequence 14578, A
C 39	32.6	6.5	474	4 US-09-621-976-18033	Sequence 18033, A
C 40	32.6	6.5	601	4 US-09-949-016-78977	Sequence 78977, A
C 41	32.6	6.5	601	4 US-09-949-016-78978	Sequence 78978, A
C 42	32.6	6.5	2108	4 US-09-270-767-12398	Sequence 12398, A
C 43	32.6	6.5	132871	4 US-09-949-016-13863	Sequence 13863, A
C 44	32.6	6.5	161607	4 US-09-949-016-12210	Sequence 12210, A
C 45	32.6	6.5	784019	4 US-09-949-016-14033	Sequence 14033, A

ALIGNMENTS

RESULT 1
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.

TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS V6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: P186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 7.8%; Score 38.8; DB 4; Length 1830121;

Best Local Similarity 49.3%; Pred. No. 2.2; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

262 ACACGACGATGATTTCTTATTTCTTATGCTGCTGACCTGCAACCGTTAT 321

Db	1506177	AAAGTAGAGTTATTTAAAGCCTTGATTATTTATGCGTGAAGAAATTACAGCAAGAAAGTTAT	1506118
Qy	322	GCCITGCGTACGCGTCTGTCTAAAGTTCTACATTTGCGCAAGTTAGCGCCAACAATTGAT	381
Db	1506117	AGCTTGATGCACTAATAATATACGAATTGTGCTTAGCATTAAGCGAAGAAAGCAATTTGCG	1506058
Qy	382	AATAATGATCCTAGATGGTTANAGCAGATGCTGTTAAGTTGGTGATACTAGATCAACT	441
Db	1506057	AGTTATGTGTTAAAGGCTTCAGGGAATGGAAGGCCATTAAGCCTGTGCTAGAAAGCAACT	1505998
Qy	442	TTAATATTTTTCACAGTTGAT	464
Db	1505997	TTAAACCTTTTGAAGATAGAT	1505975

```

1      RESULT 2
2      US-09-643-990A-1/c
3      : Sequence 1, Application US/09643990A
4      : Patent No. 6528289
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Robert D. Fleischmann
8      : Mark D. Adams
9      : Owen White
10     : Hamilton O. Smith
11     : J. Craig Venter
12     :
13     : TITLE OF INVENTION: The Nucleotide sequence of
14     : the Haemophilus influenzae Rd Genome, Fragments
15     : Thereof, and Uses Thereof
16     :
17     : NUMBER OF SEQUENCES: 1
18     : CORRESPONDENCE ADDRESS:
19     : ADDRESSEE: Human Genome Sciences, Inc.
20     : STREET: 9410 Key West Avenue
21     : CITY: Rockville,
22     : STATE: MD
23     : COUNTRY: USA
24     : ZIP: 20850
25     :
26     : COMPUTER READABLE FORM:
27     : MEDIUM TYPE: 3 1/2 inch diskette
28     : COMPUTER: Dell Pentium
29     : OPERATING SYSTEM: MS DOS v6.22
30     : SOFTWARE: ASCII Text
31     :
32     : CURRENT APPLICATION DATA:
33     : APPLICATION NUMBER: US/09/643,990A
34     : FILING DATE: 23-Aug-2000
35     : CLASSIFICATION: <Unknown>
36     :
37     : PRIOR APPLICATION DATA:
38     : APPLICATION NUMBER: 08/487,429
39     : FILING DATE: 1995-06-07
40     : APPLICATION NUMBER: 08/426,787
41     : FILING DATE: 1995-04-21
42     :
43     : ATTORNEY/AGENT INFORMATION:
44     : NAME: Kenley K. Hoover
45     : REGISTRATION NUMBER: 40,302
46     : REFERENCE/DOCKET NUMBER: PB186P1C1
47     :
48     : TELECOMMUNICATION INFORMATION:
49     : TELEPHONE: 301-610-5790
50     : TELEFAX: 310-309-8439
51     :
52     : INFORMATION FOR SEQ ID NO: 1:
53     :
54     : SEQUENCE CHARACTERISTICS:
55     : LENGTH: 1830121 base pairs
56     : TYPE: nucleic acid
57     : STRANDEDNESS: double
58     : TOPOLOGY: linear
59     :
60     : SEQUENCE DESCRIPTION: SEQ ID NO: 1:
61     :
62     :
63     :
64     :
65     :
66     :
67     :
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :
625    :
626    :
627    :
628    :
629
```

Query Match	7.8%	Score 38.8	DB 4	Length 1830121
Best Local Similarity	49.3%	Pred. NO. 2.2		
Matches 100	Conservative	0	Mismatches 103	Indels 0
Gaps				
262 ACACCAGACGATTAAGTTTCTTATTTCTCTGATGCTCGAGCCGCAACCGGTTAT	321			

Db	1506117	AAAGTAGAGTTTATTAAGCCTTTGATTATGTCGTGAGAAATTAACAGCAAGAAGTTAT	1506118
Qy	322	GCGTTGGCTAGCGCTCTGTCTAAAGTTTACATTTCCGCAACGTTGAGCCCAACAATTGAT	381
Db	1506117	AGCTTGTGCAATTAATTTAAACGAATTCGTGATTCAGCATTAAGCGAAGAAAGCATTTGCG	1506058
Qy	382	AATATATATCCTAGTGTGTANGACATGCGGTAACTGGGCGATACCTAGATCACT	441
Db	1506057	AGTTTATGTGTAAAGGCTTGAGGGGAATGGAAGGGCGATTAAGCCTGTCTTGAAAGCAACT	1505998
Qy	442	TTAATATTTTTCGAACGTGAT	464
Db	1505997	TTAAAACTTTTGAAGATAAGAT	1505975

```

1      RESULT 3
2      US-09-949-016-126480
3      : Sequence 126480: Application US/09949016
4      : Patent No. 6812339
5      : GENERAL INFORMATION:
6      : APPLICANT: VENTER, J. Craig et al.
7      : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      : TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      : FILE REFERENCE: C001307
10     : CURRENT APPLICATION NUMBER: US/09/949,016
11     : CURRENT FILING DATE: 2000-04-14
12     : PRIOR APPLICATION NUMBER: 60/241,755
13     : PRIOR FILING DATE: 2000-10-20
14     : PRIOR APPLICATION NUMBER: 60/237,768
15     : PRIOR FILING DATE: 2000-10-03
16     : PRIOR APPLICATION NUMBER: 60/231,498
17     : PRIOR FILING DATE: 2000-09-08
18     : NUMBER OF SEQ ID NOS: 207012
19     : SOFTWARE: FASTSEQ for Windows Version 4.0
20     : SEQ ID NO 126480
21     : LENGTH: 601
22     : TYPE: DNA
23     : ORGANISM: Human
24     : US-09-949-016-126480

```

Query Match	7.7%	Score 38.4	DB 4	Length 601
Best Local Similarity	60.0%	Pred. No. 0.047		
Matches	63	Conservative	0	Mismatches 42
				Indels 0
				Gaps 0
QY	374	CAATTGATTAATTAATGATCTCTGATAGTGTANAGCATGGTGGTAAAGTTGGTGATACACAG	433	
DB	422	CAATTGATTAAGAAAGATACCTGGTAGATGACCTTAATGCTTTTTCGATATGAAAATTAATTA	481	
QY	434	GATCAACTTAATATATTTTTCGAACGTTGATGTGATGTTTCTTA	478	
DB	482	GAACAGCTTAACCTATTTTAAACCTTTATTTAGATGTTTATCTA	526	

```

1 RESULT 4
2 US-09-949-016-15339
3 Sequence 15319, Application US/09949016
4 Patent No. 6812339
5 GENERAL INFORMATION:
6 APPLICANT: VENTER, J. Craig et al.
7 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9 FILE REFERENCE: CL001307
10 CURRENT APPLICATION NUMBER: US/09/949,016
11 PRIOR FILING DATE: 2000-04-14
12 PRIOR APPLICATION NUMBER: 60/241,755
13 PRIOR FILING DATE: 2000-10-20
14 PRIOR APPLICATION NUMBER: 60/237,768
15 PRIOR FILING DATE: 2000-10-03
16 PRIOR APPLICATION NUMBER: 60/231,498
17 PRIOR FILING DATE: 2000-09-08
18 NUMBER OF SEQ ID NOS: 207012
19 SOFTWARE: FastSeq for Windows Version 4.0
20 SEQ ID NO: 15339

```


RESULT 10
US-09-400-348-2
Sequence 2, Application US/09400348
Patent No. 6355778
GENERAL INFORMATION:
APPLICANT: Eckert, Joseph
APPLICANT: Alonso, Jose
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
TITLE OF INVENTION: AND PATHOGENS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6355778-rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/400,348
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/819,288
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: UPN-2949
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4746 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHEITICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 584..4468
US-09-400-348-2

Query Match 7.5%; Score 37.4; DB 3; Length 4746;
Best Local Similarity 58.6%; Pred. No. 0.29;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 TTCTTATAAACAAGAGAGAAACTTGTGCCACACTGATGGCGAAGAGAGTGAATGCG 133
Db 2391 TTGTTACATGAGAAATTAACGCAAGTTTATTGAAAGAGAGTGAAGGGGTTTCATGGG 2450

Qy 134 AAACAGAGAAGAGGAGCATGCGACCTCAATATATATTGTCGCCAAT 184
Db 2451 AAACAGAGAAGAGTCAACAAAGCTGCTCTTACAGCAACTTACTGTGGAT 2501

RESULT 11
US-08-261-822A-2
Sequence 2, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Eckert, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
TITLE OF INVENTION: and Pathogens

NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553-rls
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4747 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEITICAL: NO
ANTI-SENSE: NO
US-08-261-822A-2

Query Match 7.5%; Score 37.4; DB 1; Length 4747;
Best Local Similarity 58.6%; Pred. No. 0.29;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 TTCTTATAAACAAGAGAGAAACTTGTGCCACACTGATGGCGAAGAGAGTGAATGCG 133
Db 2391 TTGTTACATGAGAAATTAACGCAAGTTTATTGAAAGAGAGTGAAGGGGTTTCATGGG 2450

Qy 134 AAACAGAGAAGGAGCATGCGACCTCAATATATATTGTCGCCAAT 184
Db 2451 AAACAGAGAAGAGTCAACAAAGCTGCTCTTACAGCAACTTACTGTGGAT 2501

RESULT 12
PCT-US95-07744A-2
Sequence 2, Application PC/TUS9507744A
GENERAL INFORMATION:
APPLICANT: Trustees of The University of Pennsylvania
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
TITLE OF INVENTION: and Pathogens
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07744A
FILING DATE: 15-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,822

Patent No. 5955652
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph
APPLICANT: Alonso, Jose
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
TITLE OF INVENTION: AND PATHOGENS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652-rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,288
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: UPN-2949
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (Genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-819-288-1

Query Match 7.5%; Score 37.4; DB 2; Length 6172;
Best Local Similarity 58.6%; Pred. No. 0.33;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 74 TTCTTATAACAGAGGAAACTTGTGCCACACTGATGGCGAAGAGTTGATGCG 133
DB 3740 TTGTTACATGAGAGATTAACGCAAGTTATTGAAAGAGTTGAAGGGGTTTCATGG 3799
QY 134 AAACAGAGAGAGGAGCATGCACTCAATAATATCTATTGTCGCCAAT 184
DB 3800 AAACAGAGAGAGCTACCAAGAGCTCTCTACAGCAACTTACTGTGCGAT 3850

Search completed: February 26, 2005, 20:48:31
Job time: 158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 26, 2005, 19:05:25 ; Search time 499 Seconds
(Without alignments)
5913.585 Million cell updates/sec

Title: US-10-621-901-26
Perfect score: 498
Sequence: 1 ggcctcgagcaagatgcctc.....aaattctgcgaagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	17	US-10-621-901-26
2	478.6	96.1	498	10	US-09-991-936-381
3	478.6	96.1	498	17	US-10-621-901-426
4	41.8	8.4	2172	15	US-10-154-419-9
5	41.8	8.4	2634	14	US-10-154-419-7
6	40.2	8.1	8866	15	US-10-239-676-140
7	40.2	8.1	8866	15	US-10-240-453-154
8	39.6	8.0	5862	15	US-10-111-455-518
9	38.8	7.8	1830121	17	US-10-329-670-1
10	38.8	7.8	1830121	18	US-10-158-865-1
11	38.4	7.7	15295	9	US-09-764-877-3404

12	38.4	7.7	15295	17	US-10-242-515-3404	Sequence 3404, Ap
13	37.6	7.6	3020	16	US-10-252-157-347	Sequence 347, App
14	37.4	7.5	443	9	US-09-770-444-856	Sequence 856, App
15	37.4	7.5	892	13	US-10-027-633-163122	Sequence 163122,
16	37.4	7.5	892	13	US-10-027-633-163123	Sequence 163123,
17	37.4	7.5	892	17	US-10-027-633-163122	Sequence 163122,
18	37.4	7.5	892	17	US-10-027-633-163123	Sequence 163123,
19	37.4	7.5	4746	16	US-10-385-521-10	Sequence 10, Appl
20	37.4	7.5	4746	16	US-10-385-521-10	Sequence 10, Appl
21	37.4	7.5	6022	16	US-10-385-521-11	Sequence 11, Appl
22	37.4	7.5	6022	18	US-10-602-475A-14	Sequence 14, Appl
23	37.2	7.5	2333	19	US-10-741-600-535	Sequence 535, App
24	37.2	7.5	2925	19	US-10-741-600-534	Sequence 534, App
25	37.2	7.5	2982	19	US-10-741-600-533	Sequence 533, App
26	37.2	7.5	3395	19	US-10-741-600-536	Sequence 536, App
27	36.2	7.3	952	17	US-10-424-599-20505	Sequence 20505, A
28	36.2	7.3	495269	17	US-10-398-221-8	Sequence 8, Appl1
29	36.2	7.3	3011408	17	US-10-398-221-2058	Sequence 2058, Ap
30	35.6	7.1	1253	18	US-10-425-115-100766	Sequence 100766,
31	35.6	7.1	107745	18	US-10-322-281-268	Sequence 268, App
32	35.6	7.1	300000	15	US-10-262-552-33	Sequence 33, Appl
33	35.6	7.1	300000	18	US-10-703-210-33	Sequence 33, Appl
34	35.6	7.1	684707	17	US-10-398-221-9	Sequence 9, Appl1
35	35.4	7.1	269223	13	US-10-672-787-41	Sequence 41, Appl
36	34.8	7.0	798	13	US-10-027-633-124643	Sequence 124643,
37	34.8	7.0	798	17	US-10-027-633-124643	Sequence 124643,
38	34.8	7.0	200000	18	US-10-672-764A-11	Sequence 11, Appl
39	34.6	6.9	270	9	US-09-923-876-5308	Sequence 5308, Ap
40	34.6	6.9	270	10	US-09-923-876-5308	Sequence 5308, Ap
41	34.6	6.9	283	9	US-09-923-876-5178	Sequence 5178, Ap
42	34.6	6.9	283	10	US-09-923-876-5178	Sequence 5178, Ap
43	34.6	6.9	266	17	US-10-425-114-5475	Sequence 5475, Ap
44	34.6	6.9	3618	17	US-10-282-122A-35574	Sequence 35574, A
45	34.6	6.9	168575	16	US-10-178-194-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-621-901-26
; Sequence 26, Application US/10621901
; Publication No. US20040067516A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-8-C3
; CURRENT APPLICATION NUMBER: US/10/621, 901
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/319,414
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 2313
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n = unknown at position 403
US-10-621-901-26

Query Match 99.8%; Score 497; DB 17; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.6e-135;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ggcctcgagcaagatgcctccttcgactcaccacggaatgagactgtcttaccgataaatt 60
|||||

Db 1 GGGCTTGAGACAGATGCGCTTTCAGCTCAACCAAGATATGAGCTGTCTACATCAAAAT 60
Qy 61 GCCACCATTTGAATGTTCTTAATAAACAAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db 61 GCCACCATTTGAATGTTCTTAATAAACAAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Qy 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Db 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Qy 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240
Db 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240
Qy 241 GATTAACCAAAAATCTAAACAACGACGATGATTAATTTCTATTTCTTCGTATG 300
Db 241 GATTAACCAAAAATCTAAACAACGACGATGATTAATTTCTATTTCTTCGTATG 300
Qy 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTTCACATTTGCCA 360
Db 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTTCACATTTGCCA 360
Qy 361 AGCTTACGCGCAACAATTGATTAATATGATCTAGATGTTAGAGACATGTTAGT 420
Db 361 AGCTTACGCGCAACAATTGATTAATATGATCTAGATGTTAGAGACATGTTAGT 420
Qy 421 TGGGTATCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Db 421 TGGGTATCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Qy 481 ATTCTGCCAAGACTGCA 498
Db 481 ATTCTGCCAAGACTGCA 498

RESULT 2
US-09-991-936-381
; Sequence 381, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Schluchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 381
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-381

Query Match 96.1%; Score 478.6; DB 10; Length 498;
Best Local Similarity 98.8%; Pred. No. 8,9e-130;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GGGCTTGAGACAGATGCGCTTTCAGCTCAACCAAGATATGAGCTGTCTACATCAAAAT 60
Db 1 GGGCTTGAGACAGATGCGCTTTCAGCTCAACCAAGATATGAGCTGTCTACATCAAAAT 60
Qy 61 GCCACCATTTGAATGTTCTTAATAAACAAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db 61 GCCACCATTTGAATGTTCTTAATAAACAAGAGAGAAAATTGTGCCACACTGATGGCGAA 120

Qy 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Db 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Qy 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240
Db 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240
Qy 241 GATTAACCAAAAATCTAAACAACGACGATGATTAATTTCTATTTCTTCGTATG 300
Db 241 GATTAACCAAAAATCTAAACAACGACGATGATTAATTTCTATTTCTTCGTATG 300
Qy 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTTCACATTTGCCA 360
Db 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTTCACATTTGCCA 360
Qy 361 AGCTTACGCGCAACAATTGATTAATATGATCTAGATGTTAGAGACATGTTAGT 420
Db 361 AGCTTACGCGCAACAATTGATTAATATGATCTAGATGTTAGAGACATGTTAGT 420
Qy 421 TGGGTATCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Db 421 TGGGTATCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Qy 481 ATTCTGCCAAGACTGCA 498
Db 481 A-TCTGCCAAGACTGCA 497

RESULT 3
US-10-621-901-426
; Sequence 426, Application US/10621901
; Publication No. US20040067516A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Schluchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-8-C3
; CURRENT APPLICATION NUMBER: US/10/621,901
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/319,414
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 2313
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 426
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-10-621-901-426

Query Match 96.1%; Score 478.6; DB 17; Length 498;
Best Local Similarity 98.8%; Pred. No. 8,9e-130;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GGGCTTGAGACAGATGCGCTTTCAGCTCAACCAAGATATGAGCTGTCTACATCAAAAT 60
Db 1 GGGCTTGAGACAGATGCGCTTTCAGCTCAACCAAGATATGAGCTGTCTACATCAAAAT 60
Qy 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Db 121 GAGGTTGAATGCGAACAAGAGAGAGATGCGACCTCAAAATATATCTATTGCTGCC 180
Qy 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240
Db 181 CAATTAATTTCTGAGATCGGTGATCTTTATCTACACTTTAGGTATCTATATGAGAC 240

QY	241	GATTAACACAAAAAACTTAAACACAGACATGATAGTTTTCTTATTTCTCGATG	3 000
Db	241	GATTAACACAAAAAACTTAAACACAGACATGATAGTTTTCTTATTTCTCGATG	3 000
QY	301	CTCGACCTGCAACCGGTTATGCTTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGGCA	3 660
Db	301	CTCGGACCTGCAACCGGTTATGCTTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGGCA	3 660
QY	361	ACGTTACGCGCAACATTTGATATATATATGATCTTAAATGTTTANGAGCATGGTGGTAAAT	4 200
Db	361	ACGTTACGCGCAACATTTGATATATATATGATCTTAAATGTTTANGAGCATGGTGGTAAAT	4 200
QY	421	TGGGTGATCTAGAGTCAACTTTAATATTTTTTGGCAAGTTGATGATTTCTCTAAA	4 800
Db	421	TGGGTGATCTAGAGTCAACTTTAATATTTTTTGGCAAGTTGATGATTTCTCTAAA	4 800
QY	481	ATTCTGCGAAGAGCTGCA	4 980
Db	481	ATTCTGCGAAGAGCTGCA	4 980
QY	497	A-TCTGCGAAGAGCTGCA	4 970
Db	497	A-TCTGCGAAGAGCTGCA	4 970

```

RESULT 4
US-10-154-419-9
: Sequence 9, Application US/10154419
: Publication No. US20030143675A1
: GENERAL INFORMATION:
: APPLICANT: Curtis, Rory A.J.
: APPLICANT: Glucksmann, Maria Alexandra
: APPLICANT: Meyers, Rachel E.
: TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
: TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
: TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
: TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
: FILE REFERENCE: MNI-249
: CURRENT APPLICATION NUMBER: US/10/154,419
: CURRENT FILING DATE: 2002-05-22
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 99
: SOFTWARE: FastSeq Version 4.0
: SEQ ID NO 9
: LENGTH: 2172
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(2172)
: OS-10-154-419-9

```

Query Match	8.4%	Score 41.8	DB 15	Length 2112
Best Local Similarity	47.8%	Pred. No. 0.26		
Matches 155	Conservative 0	Mismatches 163	Indels 6	Gaps 1
Qy	181	CAATTAATTTCTGGAGTGGTGATCTTTAATACTACCTTAGGTGATCTCTATATGAC	240	
Db	706	CACTATTGCTGGGGGCAAGAGAACTCCTTTATATCTGGGAACGCTTCTTGAT	765	
Qy	241	GATTAACAACAAAAATCTAATAACAACGACACTGATTAAGTTTTCTAATTTCTTCGTATG	300	
Db	766	GATTCGTGACCAACACAAAGTCCTCTCTATATATAGAACCGGTTATGTCATGCATC	825	
Qy	301	CTGGACCTGCACACCGGTTATGCTTGGCTAGCGTCTCTTAAGTTCTACATTTGGCA	360	
Db	826	TTAGGCCCTGTATTTGGCTATGTATTGGGAAGACAACGTCTTAACCATTAATTAATGATTT	885	
Qy	361	ACGTGAACGCACAATATGA-----TAATATGATCTAGATGGTTANGACATGATG	414	
Db	886	GCTATGGAGAAAGCACTGATGTCACTAGAGATGATCCGCGATGGTTGGAGCTTGGTGG	945	
Qy	415	TTAAGTTGGTGATCTAGATCAACTTTAATATTTTGGCAACGTTGATGGATGTTT	474	
Db	946	ATTGGTTTCTTCATCAATGATCTTTCGTTCTTTAATAATACCTTTTCTTGCTTT	1005	
Qy	475	CCTAAATTTCTGCCAAGACTGCA	498	

Db 1006 CCAAACATTACCAGGTACAGCA 1029

```

RESULT 5
US-10-154-419-7
; Sequence 7, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL SEQUA
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44481, 67084FL, 67084ALT, FBH58295FL, 57255
; TITLE OF INVENTION: AND 57255ALT MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; PRIORITY FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 7
; LENGTH: 2634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (104)...(2275)
US-10-154-419-7

```

Query Match	8.4%: Score 41.8; DB 15; Length 2634;
Best Local Similarity	47.8%: Pred. No. 0.29;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;	
Qy	181 CAATTAATTTCTGAGTCGGTGTGATCTTTATPACTACCTTAGGTGATCCTATATGAGC 240
Db	809 CAACTATATGCTGGGGGAGGAGGAACTCCTTTATPACTCTGGAAACAGCTTCTTGAT 868
Qy	241 GATAACACCAAAAAATCTTAAAAACACGACACTGATAAGTTTTCCTATTTTCTTCGATG 300
Db	869 GATTCGTGTGCCACACACCAAGTCTTCTCTCATATPAGAACCCGGTTATGCTATGTCAATC 928
Qy	301 CTGGACCTGCAACCGGTTATGCTTGGCTAGAGCTGTCTTAAAGTTCTACATTTCGCCA 360
Db	929 TTAGGCCCTGCTATATGGCTATGTAATGGGAGACCACTGCTTAAACCATATGATTTGAT 988
Qy	361 ACCTTGAACGCCAACATTTGA-----TAAATATGATCCTAGATGTTANGAGATGGTG 414
Db	989 GCTATGGGAGAAACACTGATATGTCACGAGGATATCCGGATGTTGGAGCTTGGTGG 1048
Qy	415 TTAAAGTTGGGTGATPACTAGATCAACTTTAATATTTTGGACAGTTGATGATGTT 474
Db	1049 ATTGGGTTTCTTCATATATGATCTTTGGCTTGTCTTAAATATACCTTTTCTTGCTTT 1108
Qy	475 CCTAAATATCTGCCAGAGCTGCA 498
Db	1109 CCAAAACATTTACCGAGTACAGCA 1132

```

RESULT 6
US-10-239-676-140
; Sequence 140, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 503.1003
; CURRENT APPLICATION NUMBER: US/10/239.676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968

```

```

? DE 10019058.8
? DE 10019173.8
? DE 10032529.7
? DE 10043826.1
? PRIOR FILING DATE: 2001-04-06
? 2000-04-06
? 2000-04-07
? 2000-06-30
? 2000-09-01
? NUMBER OF SEQ ID NOS: 228
? SEQ ID NO 140
? LENGTH: 8866
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (3266, 3272, 3300, 3310, 3314..3315, 3334..3335, 3371..3372)
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (3405, 3418)
? OS-10-239-676-140

```

Query Match	8.1%;	Score 40.2;	DB 14;	length 8866;
Best Local Similarity	49.1%;	Pred. No. 1.6;		
Matches 105; Conservative	0;	Mismatches 109;	Indels 0;	Gaps 0;

Oy 273 GAAAGATTTTCTAATTCTCGATGCTGGACCTGCACCAGGTATGCCCTTGCGCTAG 332
|||
Db 4280 GAATTCGTTTAGTGAATTTGATTTTTGGCGCTTAATTAGAGCAATTCGAAGCTTTGG 4339

OY 333 CGCTGTCTAAAGTTCATCATTTGCGCAACGTTAAGCACAATTGATAATATGATCC 392
| | | | |
Db 4340 TTATGGTAAAAGTTTATGCATTTTCGTTATTTTAGTAGTGTTTTTATATTAATAGATTT 4399

QY	393	TAGATGGTTANGACA	TGCTGGTTAAGTTGGGTGATAC	TACAGATCAACTTTAATATTTT	452
b					
4400	TGACGTAATTTTGGAT	TGCTGGTAAGTTTCTTTTA	TACGTTTTTTTAAATTTTTT	4458	

QY	453	TCGACGCTGATTGGATTGTTTCTCTAAATTCG	486
Qy	453	TCGACGCTGATTGGATTGTTTCTCTAAATTCG	486
b	4460	TCGAATTAGTATTAGTTGATTATTACGGTTG	4493

```

US-10-240-453-154
; Sequence 154, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 154
; LENGTH: 8866

```

```

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (3266, 3272, 3300, 3310, 3314..3315, 3354..3355, 3371..3372)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (3405, 3418)
:
: US-10-240-453-154

```

Query Match	8.1%;	Score 40.2;	DB 15;	Length 8866;
Best Local Similarity	49.1%;	Pred. No. 1.6;		
Matches 105; Conservative	0;	Mismatches 109;	Indels 0;	Gaps 0;

Oy 273 GATTAAGTTTCTTATTTCTTCGTATCTCGACCTCGACCCGTTATGCTTGGCTAG 332
 Db 4280 GATTTCGTTACGTATGTTATTTTGGGCTTTATTAGAGCTAATCGAGGTTGG 4333

Oy 333 CGCTGTCTAAAGTTCACATTTTCGCCAAGTGAACGCCAACAAATGATATATATGATCC 392
 Db 4340 TTAATGCTAAAGATTATGCAATTTTCGTATTTTGAAGTGTGTTTAAATTAATTAAGATTT 4399

Qy	Db
393	452
TAGATGGTANAGCATGCTGGTAAAGTTGGGTCATACAGATCAACTTTAATTTT	TTTATTTT
4400	4455
TGACGTAATTTGGGATGCTGGTAAGTTTTTTGTTTATACGTTTTTTTAAATTTTTT	TTTAAATTTTTT

Qy	453	TGCAACGTTGATGGATGTTTCCTAAATTCG	486
Db	4460	TGAATTAGTATTAGTTGTATTATTACGGTTG	4493

RESULT 8
US-10-311-455-518

```

; Publication No.: US20030143606A1
;
; GENERAL INFORMATION:
;
; APPLICANT: OLEK, Alexander

```

APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System
TITLE OF INVENTION: Cytosine methylation

CURRENT APPLICATION NUMBER: US/10/311,455
 CURRENT FILING DATE: 2002-12-16
 PRIOR APPLICATION NUMBER: PCT/EP01/07537

PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1

```

: PRIOR FILING DATE: 2006-05-01
: NUMBER OF SEQ ID NOS: 2424
: SEQ ID NO 518
: LENGTH: 5882

```

1156: DNA
1157: ORGANISM: Artificial Sequence
1158: FEATURE:
1159: COMMENT: INCOMPARTION: chemically treated genomic DNA (Homo sapiens)

```

;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1463, 5077
; count: 2190456707

```

Query Match	8.0%	Score 39.6;	DB 15;	Length 5882;
US-10-311-455-518				

Matches	66;	Conservative	0;	Mismatches	45;	Indels	0;	Gaps	0;
OY	376	ATTGATATAATGATCCTAGATGTTTANGAGCATGTTGTTAACTTGGTGATAC	TACGA	435					

Dd 2872 ATGTATAATATGTCGTGTTTTTTGAAGTAAAGGTTTAAATTTGTTTTGTTAGA 2931

QY 436 TCACTTAAATTTTTCACCGTTCGATTCGATTCCTAAATTCG 486
Db 2932 TTCGATTTGATTTTTCGTAATGTTTATGTTTATTAATGTGTG 2982

RESULT 9
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US2004001850A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: P186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (120038) .. (120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) .. (121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167) .. (122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) .. (122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340) .. (131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) .. (131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) .. (139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) .. (140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) .. (142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) .. (145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) .. (145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) .. (145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) .. (147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) .. (150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) .. (152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530) .. (152530)

Query Match 7.8%; Score 38.8; DB 17; Length 1830121;
Best Local Similarity 49.3%; Pred. No. 58;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 262 ACACCGACGATGATTTTCTTATTTCTTCGTGCTCGGACCTGCAACCGGTTAT 321
DB 1506177 AAGGTAGAGTTTATTAAGCCTTGATTTATGCGTCAAGAAATTAACGCAAGAGTTAT 1506118
QY 322 GCCTTGCGTACGCGTGTCTTAAGTTCTACATTTGCGCAAGTGAAGCCCAACATGAT 381
DB 1506117 AGCTTGATGCATTAATATTAACGAATGTGCAATTAAGCGAAGAAAGCATTTGCG 1506058

QY 382 AATAATGATCCTTGATGCTTANGACATGTGTTAAGTGGATAGATCAACT 441
DB 1506057 AGTTTATGCTGTAAGCGTTGAGGGAATGGAAGGCGATTAAGCCTGTGTTAAGCAACT 1505998
QY 442 TTAATATTTTTCGCAAGTTGAT 464
DB 1505997 TTAACCTTTTGAAGATAAGAT 1505975

RESULT 10

US-10-158-865-1/c
Sequence 1, Application US/10158865
Publication No. US20040203093A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragment
FILE REFERENCE: PB166P2C1D1
CURRENT APPLICATION NUMBER: US/10/158, 865
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 09/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747) .. (4747)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921) .. (9921)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150) .. (10150)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298) .. (29298)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543) .. (36543)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551) .. (36551)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636) .. (36636)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810) .. (40810)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416) .. (44416)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905) .. (44905)
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (44975)..(44975)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (45593)..(45593)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (45732)..(45732)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (47036)..(47036)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (51334)..(51334)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (51602)..(51602)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (51786)..(51786)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (51805)..(51805)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (55369)..(55369)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (65309)..(65309)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (65313)..(65313)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (80024)..(80024)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (100691)..(100691)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (102696)..(102696)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (105121)..(105121)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (107248)..(107248)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (117136)..(117136)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (119750)..(119750)	
OTHER INFORMATION: n equals a,t,c, or g	
FEATURE:	

NAME/KEY: misc_feature	LOCATION: (119924) ..(119924)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (120038) ..(120038)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (121344) ..(121344)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (122167) ..(122167)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (122336) ..(122336)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (131340) ..(131340)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (139510) ..(139510)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (140398) ..(140398)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (142750) ..(142750)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (145058) ..(145058)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (145711) ..(145711)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (152500) ..(152500)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (152530) ..(152530)	OTHER INFORMATION: n equals a,t,c, or g

Query Match	7.8%	Score 38.9	DB 18	Length 1830121
Best Local Similarity	49.3%	Pred. No. 58		
Matches 100	Conservative 0	Mismatches 103	Indels 0	Gaps 0

OY	262	ACACCACACGATGATTTTCTATTATTTTTCGTAGTCGGACCGCAACCGCTAT	321

Db 1506177 AAGTGAAGTTTATTTAAAGCTTTGATTTATGCGTGAAGAAATTACAGCAAGAGTTAT 1506118
Qy 322 GCGTGGCTAGCGCTGTCTTAAAGTTCTACATTTCCGCAAGTTCAGCCCAAGATTGAT 381
Db 1506117 AGCTTGATGATTAATTAATTAAGATTTGATTTGATTAAGGAAAGAAAGATTGGC 1506058
Qy 382 AATTAATGATCTTGATGATTGATTANAGCATGGTGTAAAGTTGGGTGATTAAGATCACT 441
Db 1506057 AGTTTATGTGTAAAGGTTGAGAGGATGGAAGGCGATTAAGCCTGTGCTAAGAACT 1505998
Qy 442 TTAATATTTTTCGACAGCTGAT 464
Db 1505997 TTAATACTTTTGAAGATAGAT 1505975
RESULT 11
US-09-764-877-3404
; Sequence 3404, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3404
; LENGTH: 15295
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3404
Query Match 7.7%; Score 38.4; DB 9; Length 15295;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 374 CAATTGATTAATATGATCTAGATGGTTAGAGCATGGTGAAGTTGGGTGATAGTAG 433
Db 12389 CAATTGATTAATGAAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 12448
Qy 434 GATCAACTTTAATATTTTTCGCAACGTTGATGGATGTTTCTTA 478
Db 12449 GAACAAGTTAATCACTTTTAAAGCCTTTATTAAGATGTTATCTTA 12493
RESULT 12
US-10-242-515-3404
; Sequence 3404, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3404
; LENGTH: 15295
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3404
Query Match 7.7%; Score 38.4; DB 17; Length 15295;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 374 CAATTGATTAATATGATCTAGATGGTTANAGCATGGTGAAGTTGGGTGATAGTAG 433
Db 12389 CAATTGATTAATGAAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 12448
Qy 434 GATCAACTTTAATATTTTTCGCAACGTTGATGGATGTTTCTTA 478
Db 12449 GAACAAGTTAATCACTTTTAAAGCCTTTATTAAGATGTTATCTTA 12493
RESULT 13
US-10-252-157-347
; Sequence 347, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 347
; LENGTH: 3020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 372313.6
US-10-252-157-347
Query Match 7.6%; Score 37.6; DB 16; Length 3020;
Best Local Similarity 47.6%; Pred. No. 5.3;
Matches 152; Conservative 0; Mismatches 155; Indels 12; Gaps 1;
Qy 192 TGAAGTCGGTGAATCTTTATATCACTTTAGGTGATCTTATATGAGATTAACACCA 251
Db 557 TGAATAGGGGAAACCCATCTGCTTGGGTATTTCTATATGAAGATTTTGCACAA 616
Qy 252 AATATCTAAACGACGACCTGATTAAGTTTCTTATTTTCTTGATGCTGACACTGC 311
Db 617 ATTGAAATATCTCTTTTATATTTGGGCTTGAGAAACAGAGCTATTTATTTGCTCTT 676
Qy 312 AACCGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCTACATTTCCGCAAGTTGACGCC 371
Db 677 GATTGACATTTTGTGACATCATCTGTGCAATGTTATGTTGACACAGATTTGGA 726
Qy 372 AACATTTGATATTA-----TGATCTAATGTTTANAGCATGGTGTAAAG 419
Db 737 CACAGATGATCTGATATTAATCTCCACATGACACTCGTTGGGTCGTCGATGGGTTGG 796
Qy 420 TTGGGATATCTAGATCAACTTTATTTTGTGAAGTTGATGATGTTTCTTA 479
Db 797 CTTTGTGATTTGTGACAGATTAACTGCTCACTGCAATCTTTTCTTTTGCCCA 856

QY 480 AATTCTGCCAAGAGCTGCA 498
DB 857 CACACTTCCAAAGAGGCA 875

RESULT 14

US-09-770-444-856
Sequence 856, Application US/09770444
Patent No. US20020023280A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy W.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moesener, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kricke, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2027 (PARA-016PRV)
CURRENT APPLICATION NUMBER: US/09/770,444
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,502
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ. ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO 856
LENGTH: 443
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-444-856

Query Match 7.5%; Score 37.4; DB 9; Length 443;
Best Local Similarity 58.6%; Pred. No. 2.3;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 74 TTCTTATAAAGAGAGGAACTTGGCCACAGTGGGAGAGGATGATGCG 133
DB 270 TTGTTACCATGAGAAATACGCAAGTTTATTGAAAAGATGTTGAAAGGGCTTCATGGG 329
QY 134 AAACAGAGAGAGGAGCATCGCACTCAATAATATATTGCTGCCCAAT 184
DB 330 AAACAGAGAGGAGCTACCAAGAGCTCTCTCAAGCACTTACTGTGGAT 380

RESULT 15

US-10-027-632-163122/C
Sequence 163122, Application US/10027632
Publication No. US2002019837A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ. ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO 163122
LENGTH: 892
TYPE: DNA
ORGANISM: Human
US-10-027-632-163122

Query Match 7.5%; Score 37.4; DB 13; Length 892;
Best Local Similarity 50.6%; Pred. No. 3.3;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 319 TATGCTTGCTAGCGCTGTCTAAAGTTCTACATTTGCCAAGTTCAGCCACCAAT 378
DB 736 TAGGTATGCTTCTTATTCTAATTTGTAATTCAGAGGAAATCAACAAATGTC 677
QY 379 GATTAATGATCTGATGATGTTANGAGCATGCTGTTAAGTTGGTGATAGATCA 438
DB 676 TAAATATATACGATATGCGTATGACACCAAAATTAATGTTGAAATTTCTAATGTC 617
QY 439 ACTTATATTTTGGCAAGTGTGATGATGTTTCTTAAATTTCTGCCAAGGC 494
DB 616 AGTAAATCATATGCAAGCTGATTTTCTGATGATATATTCACAAATCTGC 561

Search completed: February 26, 2005, 20:57:04
Job time: 510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 18:45:12 ; Search time 3109 Seconds

(without alignments)
6097.139 Million cell updates/sec

Title: US-10-621-901-26

Perfect score: 498

Sequence: 1 ggcctcgagcaagatgcctc.....aaatctgcgaagagctgca 498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	2	BF731833 3085-57 h
2	478.6	96.1	498	4	BM056367 2092-22 h
3	147.8	29.7	993	3	CNS08HM7
4	109.2	21.9	920	8	CC131868 NDL.93017
5	107.6	21.6	398	5	BK553516 BK553516
6	73.8	14.8	444	7	CO340715 EP11205.3
7	72	14.5	412	4	BT513920 BT160014A
8	71.2	14.3	371	4	BT513842 BT160014A
9	69.2	13.9	632	5	BM618196 17006592
10	66	13.3	561	5	BK551134 BK551134
11	50.8	10.2	463	1	AU209481 AU209481
12	46.6	9.4	605	5	BU214734 603755346
13	43	8.6	484	1	AU202605 AU202605
14	42.8	8.6	840	9	CL865155 t2c08be.f
15	42.2	8.5	555	7	CO715174 DG14-245n
16	41	8.2	707	1	AJ741950 AJ741950
17	40.8	8.2	1021	9	CNS016E1
18	40.6	8.2	802	8	BZ400007 E1NAS61TF
19	40.6	8.2	838	5	BZ393310 E1NAS61TR
20	40.4	8.1	564	5	BK517949 BK517949
21	40.4	8.1	693	1	A1789248 uk52907.y
22	39.8	8.0	599	9	CPA56299 CRYC08PO
23	39.8	8.0	1101	9	CNS00782
24	39.6	8.0	329	7	CK986460 4088369 B

25	39.6	8.0	521	1	AA043744
26	39.6	8.0	725	1	AJ741939
27	39.6	8.0	760	5	BU199773 604157396
28	39.6	8.0	1101	9	CNS00EPM
29	39.4	7.9	632	9	CNS011N3
30	39.4	7.9	765	8	BZ777990
31	39.2	7.9	592	1	A1663317
32	39.2	7.9	673	1	A1788287 uk56b12.y
33	38.8	7.8	685	8	BZ023383
34	38.6	7.8	539	8	AO003681 Cpg0327B
35	38.6	7.8	654	8	CC167112 i158b07.b
36	38.6	7.8	2394	3	AK049253 Mus muscu
37	38.6	7.8	3412	3	AK049944 Mus muscu
38	38.4	7.7	645	8	AZ524922 238PBa.07
39	38.4	7.7	851	2	AW045107 um17909.Y
40	38.4	7.7	1101	9	CNS0100X
41	38.2	7.7	607	7	CN433293
42	38.2	7.7	916	9	CC971176 ZOUAF20TV
43	38	7.6	600	9	CNS07586
44	37.8	7.6	456	8	BZ616865
45	37.8	7.6	642	8	CC163691

ALIGNMENTS

RESULT 1
BF731833
LOCUS
DEFINITION
3085-57 hindgut and Malpighian tubule subcloned cDNA library
ACCESSION
BF731833
VERSION
BF731833.1 GI:22038812
KEYWORDS
EST.
SOURCE
Cenozocephalides felis (cat flea)
ORGANISM
Cenozocephalides felis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicinae;
Cenozocephalides
1 (bases 1 to 498)
REFERENCE
Gaines, P.J., Brandt, K.S., Eisele, A.M., Wagner, W.P., Bozic, C.M. and
Wlansky, N.
TITLE
Analysis of expressed sequence tags from subcloned and
unsubcloned Cenozocephalides felis hindgut and Malpighian tubule
cDNA libraries
JOURNAL
ENSEMBL
MEDLINE
PUBMED
COMMENT
12144694
Contact: Gaines, Patrick J.
Pharmaceutical Discovery Group
Heska Corporation
1613 Prospect Parkway, Fort Collins, CO 80525 USA
Tel: 970 493-7272
Fax: 970 472-1644
Email: gaines@heska.com
Seq primer: CGA GCG CCG ATC CAC TAG.
Location/Qualifiers
1. 498
/organism="Cenozocephalides felis"
/mol_type="mRNA"
/db_xref="taxon:7515"
/clone="3085-57"
/sex="Female and male (4:1 ratio)"
/tissue_type="hindgut and Malpighian tubule"
/dev_stage="unfed adult and 24 hour cat blood-fed adults
(1:1 ratio)"
/clone_lib="hindgut and Malpighian tubule subcloned cDNA
library"
/note="The library was made using the PCR-select (CDN
Subtraction Kit (CDNTECH Laboratories, Inc.) with 2 (g
hindgut and Malpighian tubule mRNA used as to make the
'tester' cDNA and 2 (g carcass (all flea tissues minus
the hindgut and Malpighian tubules) to make the 'driver'

cDNA. Suppression PCR and subtractive hybridization techniques are then used to enrich for cDNAs in the tester pool that are not also present in the driver pool.

ORIGIN

Query Match 99.8%; Score 497; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.3e-131;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGCCCTGGACAAGATGCGCTTGGACCTCACCGAATATGAGCTGTCTACGATCAAAAT 60
DB 1 GGCCCTGGACAAGATGCGCTTGGACCTCACCGAATATGAGCTGTCTACGATCAAAAT 60
QY 61 GCCACCATTTGAATCTTATATAACAGAAAGAGAAAATTGTGCCACACTGATGGCGAA 120
DB 61 GCCACCATTTGAATCTTATATAACAGAAAGAGAAAATTGTGCCACACTGATGGCGAA 120
QY 121 GGAGTTGAATGGCAACAGAAAGAGAGATGGACCTCAAAATATACTATTGTGCGCC 180
DB 121 GGAGTTGAATGGCAACAGAAAGAGAGATGGACCTCAAAATATACTATTGTGCGCC 180
QY 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATCTACACTTGAATGATCTATATGAGAC 240
DB 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATCTACACTTGAATGATCTATATGAGAC 240
QY 241 GATTAACACCAAAAATCTAAACACGACGACTGATAGTTTCTTATTTCTTGTATG 300
DB 241 GATTAACACCAAAAATCTAAACACGACGACTGATAGTTTCTTATTTCTTGTATG 300
QY 301 CTGGGACCTGGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGCGCA 360
DB 301 CTGGGACCTGGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGCGCA 360
QY 361 ACCTTACCGCAACATTTGATTAATATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
DB 361 ACCTTACCGCAACATTTGATTAATATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
QY 421 TGGGTATATCTAGATCACTTAAATATTTTGGCAACGTTGATTTGTTCCCTAAA 480
DB 421 TGGGTATATCTAGATCACTTAAATATTTTGGCAACGTTGATTTGTTCCCTAAA 480
QY 481 ATTTCGCCAAGAGCTGCA 498
DB 481 ATTTCGCCAAGAGCTGCA 498

RESULT 2
BM056367 498 bp mRNA linear EST 04-SEP-2002
LOCUS BM056367
DEFINITION 2092-22 hindgut and Malpighian tubule subtracted cDNA library
ACCESSION BM056367
VERSION BM056367.1 GI:16898176
KEYWORDS EST.
SOURCE Ctenocephalides felis (cat flea)
ORGANISM Ctenocephalides felis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicines;
Ctenocephalides.
1 (bases 1 to 498)
Gaines,P.J., Brandt,K.S., Eisele,A.M., Wagner,W.P., Bozic,C.M. and
Wisniewski,N.
Analysis of expressed sequence tags from subtracted and
unsubtracted Ctenocephalides felis hindgut and Malpighian tubule
cDNA libraries
Insect Mol. Biol. 11 (4), 299-306 (2002)

JOURNAL MEDLINE
PUBMED 12144694
COMMENT Contact: Gaines, Patrick J.
Pharmaceutical Discovery Group
Heska Corporation
1613 Prospect Parkway, Fort Collins, CO 80525 USA
```

Tel: 970 493-7272
Fax: 970 472-1644
Email: gaines@heska.com
Seq primer: CGA GCT CGG ATC CAC TAG.

FEATURES

source

1..498
/organism="Ctenocephalides felis"
/mol_type="mRNA"
/db_xref="taxon:7515"
/clone="2092-22"
/sex="female and male (4:1 ratio)"
/tissue_type="hindgut and Malpighian tubule"
/dev_stage="unfed adult and 24 hour cat blood-fed adults
(1:1 ratio)"
/clone_lib="hindgut and Malpighian tubule subtracted cDNA
library"
/note="The library was made using the PCR-select (cDNA
subtraction kit (CLONTECH Laboratories, Inc.) with 2 (g
hindgut and Malpighian tubule mRNA used as to make the
'tester' cDNA and 2 (g carcass (all flea tissues minus
the hindgut and Malpighian tubules) to make the 'driver'
cDNA. Suppression PCR and subtractive hybridization
techniques are then used to enrich for cDNAs in the
tester pool that are not also present in the driver pool."

ORIGIN

Query Match 96.1%; Score 478.6; DB 4; Length 498;
Best Local Similarity 98.8%; Pred. No. 2.5e-126;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1 GGCCCTGGACAAGATGCGCTTGGACCTCACCGAATATGAGCTGTCTACGATCAAAAT 60
DB 1 GGCCCTGGACAAGATGCGCTTGGACCTCACCGAATATGAGCTGTCTACGATCAAAAT 60
QY 61 GCCACCATTTGAATCTTATATAACAGAAAGAGAAAATTGTGCCACACTGATGGCGAA 120
DB 61 GCCACCATTTGAATCTTATATAACAGAAAGAGAAAATTGTGCCACACTGATGGCGAA 120
QY 121 GGAGTTGAATGGCAACAGAAAGAGAGATGGACCTCAAAATATACTATTGTGCGCC 180
DB 121 GGAGTTGAATGGCAACAGAAAGAGAGATGGACCTCAAAATATACTATTGTGCGCC 180
QY 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATCTACACTTGAATGATCTATATGAGAC 240
DB 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATCTACACTTGAATGATCTATATGAGAC 240
QY 241 GATTAACACCAAAAATCTAAACACGACGCTGATAGTTTCTTATTTCTTGTATG 300
DB 241 GATTAACACCAAAAATCTAAACACGACGCTGATAGTTTCTTATTTCTTGTATG 300
QY 301 CTGGGACCTGGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGCGCA 360
DB 301 CTGGGACCTGGCAACCGGTTATGCTTGGCTAGCGTCTGTCTAAAGTTCTACATTTGCGCA 360
QY 361 ACCTTACCGCAACATTTGATTAATATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
DB 361 ACCTTACCGCAACATTTGATTAATATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
QY 421 TGGGTATATCTAGATCACTTAAATATTTTGGCAACGTTGATTTGTTCCCTAAA 480
DB 421 TGGGTATATCTAGATCACTTAAATATTTTGGCAACGTTGATTTGTTCCCTAAA 480
QY 481 ATTTCGCCAAGAGCTGCA 498
DB 481 ATTTCGCCAAGAGCTGCA 497
```

RESULT 3
CNS08HM7 993 bp mRNA linear HTC 07-JAN-2003
LOCUS CNS08HM7
DEFINITION Single read from an extremity of a full-length cDNA clone made from
Anopheles gambiae total adult females. 5-PRIME end of clone

FK0AA19CA02 of strain 6-9 of *Anopheles gambiae* (African malaria mosquito).
 BX012251
 VERSION BX012251.1 GI:27561471
 KEYWORDS HTc.
 SOURCE *Anopheles gambiae* (African malaria mosquito)
 ORGANISM *Anopheles gambiae*
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 Anopheles.
 1 (bases 1 to 993)
 REFERENCE
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 FEATURES
 source
 Location/Qualifiers
 1..993
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="6-9"
 /db_xref="taxon:7165"
 /clone="FK0AA19CA02"
 /plasmid="pME18S-FL"
 /note="end : 5-PRIME"

Query Match 29.7%; Score 147.8; DB 3; Length 993;
 Best Local Similarity 62.7%; Pred. No. 3.2e-31;

Matches 262; Conservative 0; Mismatches 153; Indels 3; Gaps 2;

QY 1 GGCCCTGGAACAATGCTTCGACCTGACCGAATATGAGTGTCTAGATCAAAAT 60
 DB 577 GGCCCTGGAAGAGCGCCCTCTCCCTGACCACTGAGTACGTGC-ACGTACGATGCCAAC 635
 QY 61 GCCACCATTAAGTCTTAATTAACAAGAGAGAACTTTGTGCCACACTGATGGCAA 120
 DB 636 CAGACCGCGAGTATGAGAACGCCAGAACCCAAATCTCTGCCGCAAGATGCCACT 695
 QY 121 GGAGTT-GAATCGAAGAACAGAGAGAGATCGACCTCAATTAATCTATTGCTG 178
 DB 696 CGCGGTGCGGAATGTGAACCGAGAGGACCTGCGCACCAAGATGTCTGTTATG 755
 QY 179 CCCAATTAATTTCTGAGTGGTGATCTTATATCACTTAAGTGTATCCATATG 238
 DB 756 CCCAGTTATCTCCGCAATGGAAGCTGCTCTACTACAGCTGGCGCTCGTCAATGG 815
 QY 239 ACGATAACACCAAAAATCTAAACACGACACTGATAGTTTTCTTATTTCTTGTA 298
 DB 816 ACGATAACATCAAGAGTCAAAAACCGGCCCTGTCAGTGTCTGTACTTTCTGCGA 875
 QY 299 TGGCTGGACCTGCAACCGGTTATGCTTGGCTAGCGTGTCTTAAATTCTACTTTGCG 358
 DB 876 TGGCTGGACCGGCATCGGTACACGCTGGCGTCTCTTCTGAACTGTACATCTCGC 935
 QY 359 CACGTTGAGCGCAACATGATATATATGATCTTATGTTTANGAGATGGTGT 416
 DB 936 CCTGATGAGCGCGACGATCAAGCAATCCGACCGCGGTGCTGGTCTGGTGTGAT 993

RESULT 4
 LOCUS CC131868 920 bp DNA linear GSS 16-APR-2003
 DEFINITION ND1_93017.T7 Notre Dame Liverpool Aedes aegypti genomic clone
 ACCESSION CC131868
 VERSION CC131868.1 GI:30000923
 KEYWORDS GSS.
 SOURCE Aedes aegypti (yellow fever mosquito)
 ORGANISM Aedes aegypti
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes; Stegomyia.

REFERENCE 1 (bases 1 to 920)
 AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.
 JOURNAL BAC end sequencing of Aedes aegypti
 COMMENT Unpublished (2003)
 Other_GSSs: ND1_93017.SP6
 Contact: Brendan Loftus
 Department of Eukaryotic Genomics
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: entae@tigr.org
 Library was provided by David Severson
 Seq primer: T7
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1..920
 /organism="Aedes aegypti"
 /mol_type="genomic DNA"
 /strain="Liverpool1"
 /db_xref="taxon:7159"
 /clone="ND1_93017"
 /clone_1lb="Notre Dame Liverpool1"
 /note="Vector: pBCBAC1; Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 21.9%; Score 109.2; DB 8; Length 920;
 Best Local Similarity 64.2%; Pred. No. 3.9e-20;

Matches 181; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

QY 1 GGCCCTGGAACAATGCTTCGACCTGACCGAATATGAGTGTCTAGATCAAAAT 60
 DB 521 GGCCCTGGAAGAGCGCCCTCTCCCTGACCACTGAGTACGTGC-ACGTACGATGCCAAC 580
 QY 61 GCCACCATTAAGTCTTAATTAACAAGAGAGAACTTTGTGCCACACTGATGGCAA 119
 DB 581 CAGCTTAGAAGATGTTTGAAGCAGAGAACCGAAGCTTATCCGGGTACGGAAGT 640
 QY 120 -AGAGTTGAATGCGAAGAGAGAGAGATCGACCTCAATTAATCTATTGCT 177
 DB 641 CAGAGACTAATGTGAGTGAAGAGAGAGATCTAGCCCTCAGTGTGCTGTTG 700
 QY 178 GCCAATTAATTTCTGAGTGGTGATCTTATATCACTTAAGTGTATCCATATG 237
 DB 701 GCAAGTTCAATTTCCGGGTGAGAGATTCAGTATTAACACATGAGCGTATCGTACATG 760
 QY 238 GAGGATACACCAAAAATCTAAACACGACACTGATAGT 279
 DB 761 GATGACACATCAAGAAATCAAAAATCTCTGCACTGTTAGT 802

RESULT 5
 LOCUS BX553516 398 bp mRNA linear EST 10-OCT-2003
 DEFINITION BX553516 Glossina morsitans morsitans adult infected gut Glossina morsitans morsitans cDNA clone Tse129h11_q1c, mRNA sequence.
 ACCESSION BX553516
 VERSION BX553516.1 GI:33377668
 KEYWORDS EST.
 SOURCE Glossina morsitans morsitans
 ORGANISM Glossina morsitans morsitans
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Hippoboscidae; Glossinidae; Glossina.
 1 (bases 1 to 398)
 REFERENCE
 AUTHORS Lehane,M.J., Aksoy,S., Gibson,W., Kethornou,A., Berriman M., Hamilton,J., Soares,M.B., Bonaldo,M.F., Lehane,S. and Hall,N.
 TITLE Adult midgut expressed sequence tags from the tsetse fly Glossina morsitans morsitans and expression analysis of putative immune response genes

University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499

Email: generobi@life.uiuc.edu

This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation Award in Functional Genomics to G.E. Robinson and an NSF Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.

PCR Primers

FORWARD: TATATGACCTACTATAGG

BACKWARD: ATTAACTCTACTAAG

Plate: BB160014A11 row: F column: 04

Seq primer: AGCGATTAACATTTCACACAGA

High quality sequence stop: 412.

FEATURES

source

1. .412

/organism="Apis mellifera"

/mol_type="mRNA"

/strain="mixed strains of European bees, predominantly

A.m. ligustica"

/db_xref="taxon:7460"

/clone="BB160014A11F04"

/sex="female"

/tissue_type="brain"

/dev_stage="adult worker honey bee"

/lab_host="DH10B"

/clone_lib="Bee Brain Normalized Library, BB16"

/note="Organ: Brain; Vector: pTV73-Pac; Site_1: EcoRI;

Site_2: NotI; The BB16 library was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996), Genome Research 6(9): 791-806. RNA was

prepared from dissected brains of adult worker bees of

various ages and various behavioral groups."

ORIGIN

Query Match 14.5%; Score 72; DB 4; Length 412;

Best Local Similarity 57.6%; Pred. No. 1.6e-09; Mismatches 95; Indels 0; Gaps 0;

Matches 129; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Db 142 GAAGGAGCATCGACCTCAATATATCTATTTGCTGCCAATTAATTTCTGAGTCGGT 201

239 GATGTCAGATTCCTGACTGCTTGGTCTTCTCCCAATTTATCTTGGATTGGT 180

Qy 202 GGATCTTTTACTACACTTTAGGTGATCTCTATATGACGATTAACACCAAAAATCTAAA 261

Db 179 ACCACTTTTATACGCTCTTGTCAACCTACCTGATGACATACAAAGAAAGAAAT 120

Qy 262 ACACGACGACTGATTAAGTTTCTTATTTCTTGTGATGCTGGACCTGCAACGGTTAT 321

Db 119 ACTCTATAGCTTTTACGTTTACCTTGTCTGTAAGAACAGTCGAGCCAGCAATGATTT 60

Qy 322 GCCTTGCTAGCGCTGCTTAAGTTCTTACATTTGCGCAAGTT 365

Db 59 TTACTAGGTTACGCTGTCTCAGTTTGTATCATGATCCAGTTT 16

RESULT 8
LOCUS BB160014A10F04.5 Bee Brain Normalized Library, BB16 Apis mellifera
DEFINITION CDNA clone BB160014A10F04 5', mRNA sequence.
ACCESSION B1513842
VERSION B1513842.1 GI:15364207
KEYWORDS EST.
SOURCE Apis mellifera (honey bee)
ORGANISM Apis mellifera

RESULT 8
LOCUS BB160014A10F04.5 Bee Brain Normalized Library, BB16 Apis mellifera
DEFINITION CDNA clone BB160014A10F04 5', mRNA sequence.
ACCESSION B1513842
VERSION B1513842.1 GI:15364207
KEYWORDS EST.
SOURCE Apis mellifera (honey bee)
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
Aplidae; Apidae.
1 (bases 1 to 371)
Whitfield,C.W., Bonaldo,M.F., Kumar,C.G., Liu,L.,

TITLE

Pardinae,J., Robertson,H.M., Soares,B. and Robinson,G.E.

Annotated expressed sequence tags and cDNA microarrays for studies

of brain and behavior in the honey bee

Genome Res. 12 (4), 555-566 (2002)

21929762

11932240

COMMENT

Contact: Gene E. Robinson

Department of Entomology

University of Illinois

505 S. Goodwin Ave., Urbana, IL 61801, USA

Tel: 217 265 0309

Fax: 217 244 3499

Email: generobi@life.uiuc.edu

This research was funded by the University of Illinois Critical

Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation

Award in Functional Genomics to G.E. Robinson and an NSF

Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.

PCR Primers

FORWARD: TATATGACCTACTATAGG

BACKWARD: ATTAACTCTACTAAG

Plate: BB160014A10 row: F column: 04

Seq primer: AGCGATTAACATTTCACACAGA

High quality sequence stop: 371.

Location/Qualifiers

1. .371

/organism="Apis mellifera"

/mol_type="mRNA"

/strain="mixed strains of European bees, predominantly

A.m. ligustica"

/db_xref="taxon:7460"

/clone="BB160014A10F04"

/sex="female"

/tissue_type="brain"

/dev_stage="adult worker honey bee"

/lab_host="DH10B"

/clone_lib="Bee Brain Normalized Library, BB16"

/note="Organ: Brain; Vector: pTV73-Pac; Site_1: EcoRI;

Site_2: NotI; The BB16 library was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996), Genome Research 6(9): 791-806. RNA was

prepared from dissected brains of adult worker bees of

various ages and various behavioral groups."

ORIGIN

Query Match 14.3%; Score 71.2; DB 4; Length 371;

Best Local Similarity 58.5%; Pred. No. 2.6e-09; Mismatches 88; Indels 0; Gaps 0;

Matches 124; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Db 142 GAAGGAGCATCGACCTCAATATATCTATTTGCTGCCAATTAATTTCTGAGTCGGT 201

Db 216 GATGTCAGATTCCTGACTGCTTGGTCTTCTCCCAATTTATCTTGGATTGGT 157

Qy 202 GGATCTTTTACTACACTTTAGGTGATCTCTATATGACGATTAACACCAAAAATCTAAA 261

Db 156 ACCACTTTGATTAACGCTCTTGTCAACCTACCTGATGACATACAAAGAAAGAAAT 97

Qy 262 ACACGACGACTGATTAAGTTTCTTATTTCTTGTGATGCTGGACCTGCAACGGTTAT 321

Db 96 ACTCTATAGCTTTTACGTTTACCTTGTCTGTAAGAACAGTCGAGCCAGCAATGATTT 37

Qy 322 GCCTTGCTAGCGCTGCTTAAGTTCTTACATTTGCGCAAGTT 353

Db 36 TTACTAGGTTACGCTGTCTCAGTTTGTATCAT 5

RESULT 9
LOCUS BM618196
DEFINITION 17000659201984 A.Gam.ad.cDNA1 Xanthopelae gambelae cDNA clone
ACCESSION BM618196
VERSION BM618196.1 GI:18916484

RESULT 9
LOCUS BM618196
DEFINITION 17000659201984 A.Gam.ad.cDNA1 Xanthopelae gambelae cDNA clone
ACCESSION BM618196
VERSION BM618196.1 GI:18916484

KEYWORDS	EST..
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.
REFERENCE	1 (bases 1 to 632)
AUTHORS	Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE	Celera Anopheles gambiae EST project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Holt R.A. Celera Genomics 45 W. Gude Dr., Rockville, MD 20850, USA Tel.: 240453151 Fax: 2404534580 Email: HoltR@celera.com Plate: NU01004SFB row: F column: 10 Seq primer: M13 Reverse
FEATURES	Location/Qualifiers
source	1..632 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="RSP-SF (Reduced susc. to Permethrin - std. chromosome)" /db_xref="taxon:7165" /cclone="19600447009427" /dev_stage="Adult" /lab_host="DH10b" /clone_1lb="A.Gam.ad.cDNA1" /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; whole adult mosquitoes (mixed sex) frozen on liquid nitrogen, cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."
ORIGIN	
Query Match	13.9%; Score 69.2; DB 4; Length 632;
Best Local Similarity	52.7%; Pred.No.1.le-08;
Matches 149; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	
QY	194 GAGTCGGTGATCTTTTAACTACACTTAAGGTGTTCTCTATTGTGACGATAACACCAAAA 253
Db	350 GTGTGACGTGTGTGTGAATTGAGGTTCTGTGTGTGTGTGCATCCCGCCGAC 409
QY	254 AATCAAANAACAACGACATGATTAAGTTTTTCTTATTCTTGTATGCGCAGACTGCA 313
Db	410 AATGAATTCGAACACAGCTCTGTGCATACCATCTCCCTGCGCACGTTTGACCGGTGG 469
QY	314 CCGGTATAGCCTTGCTGAGCGTCTGTCAAAGTTCTACATTTGCCAACGTTGACGCCAA 373
Db	470 ATGATTTTGGCCTCGGATCACTTTGGCGCTGAAGAATCTACATGATCTCAAGAAAGACGCCA 529
QY	374 CAATTGATTAATTAATGATCTTAGATGTTANGACATGTTGTTAAGTTGGGTGATCTAG 433
Db	530 TCATTGACAGCTCCGATCCAGGTCGCTTGGAACATGTGTGCTCGCTGGAATATGCTCG 589
QY	434 GATCAACTTTAATTTTTTTCGAACGTTGATTTGATTTTCC 476
Db	590 GTGTGGCGATGATCACTTTCGCTTTTCTGTATCGGATGTTTC 632
RESULT 10	
LOCUS	BX551134
DEFINITION	BX551134 Glossina morsitans morsitans adult infected gus Glossina
ACCESSION	BX551134
VERSION	BX551134.1 GI:33374988
KEYWORDS	EST.
SOURCE	Glossina morsitans morsitans
ORGANISM	Glossina morsitans morsitans

REFERENCE	AUTHORS	TITLE
1	(bases 1 to 561)	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotia; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Hippoboscidae; Glossinidae; Glossina.
2	(bases 1 to 561)	Lehane, M.J., Aksoy, S., Gibson, M., Keshornou, A., Berriaman, M., Hamilton, J., Soares, M.B., Bonald, M.F., Lehane, S. and Hall, N. Adult midgut expressed sequence tags from the tsetse fly Glossina morsitans morsitans and expression analysis of putative immune response genes
3	Genome Biol. 4 (10), R63 (2003)	
4	14519198	
5	Contact: Hall N	
6	Pathogen Sequencing Unit	
7	The Sanger Institute The Wellcome Trust Genome Campus	
8	Hinxton, Cambridge, CB10 1SA, UK	
9	Request for clones, please contact: Mike Lehane	
10	Prof. M.J. Lehane	
11	School of Biological Sciences,	
12	University of Wales,	
13	Bangor LL57 2UW	
14	All clones with suffix q1c are reverse primer reads starting at 5' end of the cDNA all plc reads are from	
15	the 3' end.	
16	Location/Qualifiers	
17	1..561	
18	/organism="Glossina morsitans morsitans"	
19	/mol_type="mRNA"	
20	/sub_species="morsitans"	
21	/db_xref="taxon:37546"	
22	/clone="Tse116g08.q1c"	
23	/cissue_type="adult infected gut"	
24	/clone_lib="Glossina morsitans morsitans adult infected gut"	
25	/note="country: Zimbabwe; EST from adult gut infected with T.brucei"	
26	ORIGIN	
27	Query Match	13.3%; Score 66; DB 5; Length 561;
28	Best Local Similarity	59.4%; Pred. No. 9e-08;
29	Matches 111; Conservative 0; Mismatches 76; Indels 0; Gaps 0;	
30	QY	304 GGACCTGCACCGGTAAAGCTTGGCTGACGCTGCTCTAAAGTTCTACATTTCGCCACG 363
31	DB	2 GGTCCGATGATGGGCTTTTTCGCTTAATTTCTTAAATATGATTATGATCGCACT 61
32	QY	364 TTGACGCCAATGATATATATGATCCTGATGATGTATGAGCAGTGGTAAATGG 423
33	DB	62 AAAATCTCTTAATCGATATATAGATCCGCTGCTGCGAGCTGTGATGCTGGTTGG 121
34	QY	424 GTGATCTAGATCAACTTAAATATTTTTCGCAAGTGAATGATGTTCTTAAAT 483
35	DB	122 ATGATCTGGGATCATGATATATATTTCCGGTTAATAGGCTTATTCCTAAAG 181
36	QY	484 CTGCCAA 490
37	DB	182 CTACTTA 188
38	RESULT 11	
39	AU209481	463 bp mRNA linear EST 17-JUL-2001
40	LOCUS	AU209481 unpublished oligo-capped cDNA library, stage L1
41	DEFINITION	Caenorhabditis elegans cDNA clone yk1015f10 5', mRNA sequence.
42	ACCESSION	AU209481
43	VERSION	AU209481.1 GI:14845081
44	KEYWORDS	EST.
45	SOURCE	Caenorhabditis elegans
46	ORGANISM	Caenorhabditis elegans
47	REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.
48	AUTHORS	1 (bases 1 to 463)
49		Kohara, Y., Shin-1, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

TITLE and Sugano, S.
JOURNAL A complementary view of the C.elegans genome
COMMENT Unpublished (2001)
Contact: Yuiji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

source

1. 463
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK105F10"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, stage L1"

ORIGIN

Query Match 10.2%; Score 50.8; DB 1; Length 463;
Best Local Similarity 47.4%; Pred. No. 0.002;
Matches 148; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

126 TGAATCGAAACGAGAGAGAGATGCACTCAATTAATATCTATTTCTGCCAATT 185
118 GGTACTTGAGGTGGAAGAAACAAATGCAATTTCTCTTGATTAATGACGACAA 177
118 GGTACTTGAGGTGGAAGAAACAAATGCAATTTCTCTTGATTAATGACGACAA 177
246 CACCAAAATCTTAACACACACATGATGATTTCTTATTTCTTGATGCTCG 305
178 TGTCAAAAACAAATCTGCAATTAATCTTTGTTGATGATTTCTTGAAAACTCTTG 237
306 ACCTGCACCGGTATGCTTGTGCTGATGCTGCTTAAGTTCTTCAATTTGCCAACGTT 365
238 TCCGGTATTTGACCTTCTAGTGTGCTGAGGAGTTGACAAAGTTATACGTTATTTATCC 297
366 GAGCCCAACATGATTAATATGATCTAGATGTTTNGACATGCTGTTAAGTGGT 425
298 TCCACAGAGCTTACACCACTGACCAATGTGATTTGTTGCTGTGGCTGGGATTTT 357
426 GATCTAGATC 437
358 AATTTTGGAC 369

RESULT 12
LOCUS BU214734 605 bp mRNA linear EST 25-NOV-2002
DEFINITION 603755346F1 CSROCHN04 Gallus gallus cDNA clone CHEST65K12 5', mRNA
sequence.
ACCESSION BU214734
VERSION BU214734.1 GI:25391353
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE
AUTHORS Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
TITLE A Comprehensive Collection of Chicken CDNs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22355534
PUBMED 12445392

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1. 605
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, H1sex"
/db_xref="taxon:9031"
/clone="CHEST65K12"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/lab_host="DH10B"
/notes="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo (dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 9.4%; Score 46.6; DB 5; Length 605;
Best Local Similarity 49.2%; Pred. No. 0.034;
Matches 121; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

167 TACTATTGCTGCCCAATTAATTTCTGAGTGGGTGATCTTTATATACATTAGTG 226
145 TATTCATTTTGCCCAATTTCTATCGCATGGCTCCACACCATTAATACCTTGGAC 204
227 TATCCATATGAGCAATTAACCAAAAATCTAAACACGACGATTAAGTTTCTT 286
205 CAACTTACTGATGACAAATGTCAGAAAGAAACGCTGCTTACTTACCTACATGT 264
287 ATTTCTTCTGATGCTGCACTGCAACCGTTATGCTTGTGCTGCTGCTTAAGT 346
265 ATGTCATGGAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 324
347 TCTTACATTTGCCCACTTACCCCAATTTGATTAATGATCTTCTGATGTTAG 406
325 TCTATGTTGATCTTACGACCAACTGTTATTTATGACGAAATGATCTTCAATTG 384
407 CATGGT 412
385 ACTGGT 390

RESULT 13
LOCUS AU202605 484 bp mRNA linear EST 30-MAY-2003
DEFINITION AU202605 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone YK810e11 5', mRNA sequence.
ACCESSION AU202605
VERSION AU202605.2 GI:31213182
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE
AUTHORS Rhabditidae; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 484)

AUTHORS Kohara,Y., Shin-i.T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT On Jul 17, 2001 this sequence version replaced gi:14832514.

CONTACT: Yoji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..484
Location/Qualifiers

/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="Yk810e11"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"
/clone_1lb="unpublished oligo-capped cDNA library, stage L2"

ORIGIN

Query Match 8.6%; Score 43; DB 1; Length 484;
Best Local Similarity 46.8%; Pred. No. 0.35;
Matches 133; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 154 GCACCTCAATATTAATTAATTTGCTGCCCAATTAATTTGGAGTCGGTGATCTTTATAC 213
DB 1 GGACCAATTCATGATGATTTTCGAGAGACTTTTGGTACTTGGAGTTGAAACAATGCCA 60
QY 214 TACACTTGAAGTATCTATGATGAGACATACCAACCAAAATCTAAACACGACATG 273
DB 61 TTTTCTCTGGATTAACATTAATGACAGACATGTCACAAAATCTGCCATTTATAC 120
QY 274 ATAAAGTTTCTTAATTTCTTCGTATGCTGGAAGCTGCAACCGGTTATGCTTGCTAGC 333
DB 121 TTTTCGTTCATGTTCTTTGGAATCTTGATCCGGTTATGGAATTTAGTTGGGG 180
QY 334 GTCTGCTAAAGTTTCAATTTGCGCAACGTTGACGCCCAATGATTAATGATCTT 393
DB 181 CAGTTGAACAAGTTAATACGTTGATTTTAATCTCCACAGAGACTTACACCACTTGACCA 240
QY 394 AGATGTTANAGACATGAGTGAAGTGGGTGATCTAGATC 437
DB 241 ATGTGATTTGCTGCTGGTGGCTGGATTTTAAATTTTGGAAC 284

RESULT 14
CL865155 840 bp DNA linear GSS 25-AUG-2004
LOCUS t2c08be.f1 TAMBT Bos taurus genomic clone t2c08be, genomic survey
DEFINITION sequence.
ACCESSION CL865155
VERSION CL865155.1 GI:51538548
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 840)
AUTHORS Lin,S., Najjar,F.Z., Adelson,D., Gill,C.A. and Roe,B.A.
TITLE Bovine BAC End Sequences from Library TAMBT
JOURNAL Unpublished (2003)
COMMENT Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912

CONTACT: Fax: 405 325 7762
Email: broeou.edu
CLASS: BAC ends
High quality sequence start: 398
High quality sequence stop: 654.

FEATURES
source
1..840
Location/Qualifiers

/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"
/clone="t2c08be"
/sex="Male"
/cell_type="Blood"
/clone_1lb="TAMBT"
/note="Vector: pBeloBAC11, site 1: HindIII, site 2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."

ORIGIN

Query Match 8.6%; Score 42.8; DB 9; Length 840;
Best Local Similarity 50.0%; Pred. No. 0.45;
Matches 107; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 50 ACATCAAAATGCCACCATTTGAAGTCTTAATAACAGAAAGAGAAACTTGTGCCACA 109
DB 422 AGGACAGACATGCTCTCTTAATTTGTTAACTTAATTAACAGAAATCTTGGCATAAA 481
QY 110 CTGATGGCGAAGAGTGAATGCGAAACAGAAAGAGAGCATGCGACCTCAATATATAC 169
DB 482 CTAAAGCTGTTGATCACTCAAGACAAATCTTTCACATGAGATTGCACAGTTGTATGA 541
QY 170 TATTTCCTGCCAATTAATTTCTGAGTCGGTGATCTTTAATCACTTAAGTGTAT 229
DB 542 GCCTGAGACTTAATTAATTTCTTGCATCAATGAAGGGATCTCCAGGTCAAGCTTAT 601
QY 230 CCTATATGACATTAACCAAAAATCTTAAAC 263
DB 602 GGCATATATGATGGAAAGTAAAGAAATCTTACAC 635

RESULT 15
CO715174/c 555 bp mRNA linear EST 27-JUL-2004
LOCUS DG14-245n18 DG14-muscle Canis familiaris cDNA 3', mRNA sequence.
DEFINITION CO715174
ACCESSION CO715174
VERSION CO715174.1 GI:50686890
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris

REFERENCE 1 (bases 1 to 555)
AUTHORS Schlueter,T., Hermanns,J., Weindel,M., Schuete,D., Kranz,H.,
Henrich,J., and Loeberbeck,R.
TITLE Dog arrayTAG cDNA clone collection
JOURNAL Unpublished (2004)
COMMENT Contact: Thomas Schlueter
LION bioscience AG
Waldfederstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schueter@lionbioscience.com.
FEATURES
source
1..555
Location/Qualifiers

/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="DG14-muscle"

ORIGIN /note="Organ: muscle; Vector: Dog pbluescript LION"

Query Match 8.5%; Score 42.2; DB 7; Length 555;
 Best Local Similarity 48.0%; Pred. No. 0.62; Mismatches 129; Indels 0; Gaps 0;
 Matches 119; Conservative 0;

Qy	167 TACTATTGCTGCCAATTATTTCTGAGTCGGTGGATCTTTATACACTTAGATG 226
Db	338 TATTCAATTTGTGACAGATTCTCATGGAAITGGITCCACACCTATTATACCTTGGAC 279
Qy	227 TATCTATATGACGATTAACACCAAAAAATCTAAACACGACGATTAAGTTTCTT 286
Db	278 CAACCTACTTAGATGACATGTCAAGAAAGAAAGCATCTTGTACCTAGCCATCATGT 219
Qy	287 ATTTTCTTGTATGCTGGAACCGGTTATGCTTGGCTAGAGTCTGTCTAAGT 346
Db	218 ATGTCAATGGAGCACTGTGCTGCACTGGATTTATTAAGTGAATCTTATTTGATT 159
Qy	347 TCTACATTTGCGCAACGTTGACGCCCAATTTGATAAATGATCTAGATGTTANGAG 406
Db	158 TTTATGTCGATCCGAAATCTTGTAACTTGACGAATGATCTCTGTTTCATTTGAA 99
Qy	407 CATGGTG 414
Db	98 ACTGGTG 91

Search completed: February 26, 2005, 20:45:51
 Job time : 3119 secs

